DOCUMENT RESUME

ED 228 859 FL 013 630

AUTHOR Hayes, Zoe Ann
TITLE Limited Language Proficiency: A Problem in the Definition and Measurement of Bilingualism. Final Report.
INSTITUTION InterAmerica Research Associates, Rosslyn, Va.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Jan 82
CONTRACT 400-79-0042
NOTE 189p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC08 Plus Postage.
DESCRIPTORS Academic Achievement; *Bilingualism; *Bilingual Students; Evaluation Methods; *Language Proficiency; Measurement Techniques; *Measures (Individuals); *Mexican Americans; Primary Education; *Verbal Development

ABSTRACT

The phenomenon of limited bilingual language proficiency is reported to have negative academic and/or cognitive consequences, especially among Mexican-American minority language students. Where such students have been diagnosed, fears of concomitant congitive retardation are also expressed. Research was undertaken to understand the phenomenon better and to investigate its possible cognitive/academic consequences. The study defined the meaning of knowing a language, described bilingualism, and discussed bilingualism and cognitive development. The subjects for the investigation were Mexican-American third grade students in Santa Clara County (California) School districts. Analyses of all the test data used in the study indicated that no two language proficiency tests gave the same information about a child, and that the measure of communication skills was the best predictor of achievement. The results of the investigation do not lend support to the idea that limited bilingual language proficiency is widespread. Results also indicate that, among students who do exhibit this dual language limitation, no cognitive deficits are found. (AMH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED228859

Limited Language Proficiency: A Problem in the Definition and Measurement of Bilingualism

HJA

Zoe Ann Hayes

University of Nevada

FL 013 630

FINAL REPORT

January 1982

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

The research reported here was performed as part of the Assessment of Language Proficiency of Bilingual Persons (ALPBP) project. The ALPBP project was funded by the National Institute of Education (NIE Contract #400-79-0042) and administered by InterAmerica Research Associates, Inc., Rosslyn, Virginia. The opinions expressed herein are of the researchers and do not reflect those of the National Institute of Education or InterAmerican Research Associates, Inc.

© Copyright
1982
by

Zoe Ann Hayes

Precis

An assumption often voiced is that "disadvantaged" children are non-verbal or are verbally deprived. These children are said to have an enormous difficulty expressing themselves verbally in many situations. . . Consaquently, there has arisen a rather firm belief that "disadvantaged" children are basically inarticulate, that they lack the verbal ability so important in reading and eventual school success. This assumption cannot be well supported in the case of Chicano Spanish-speaking children, as we see when we analyze Chicano language characteristics and learn to appreciate the exact nature of Chicano verbal functioning.

Ernest Garcia
El Lenguaje de los Chicanos
1975

ACKNOWLEDGMENTS

This study is the result of the hard work of many people. Without their love, caring and support the research could never have been completed. I would like to express my deepest appreciation to them all.

I thank:

Dr. Robert L. Politzer for all of his time and insight offered over the years; he is a mentor in the truest sense;

Dr. Edward Haertel for advisement beyond the call of duty;

Dr. M. Beatríz Arias for being a role model and for looking between the lines;

The late Dr. Alfredo Castañeda for his guidance.

Manny, Lila, Susan, Melanie, Pearl, Marcos and Bob for all of their love and support; Anna, Sam and Herman for the wonderful memories;

InterAmerica Research Associates, the National Institute of Education, Drs. Charlene Rivera, Ellen Rosansky and Dan Ulibarri for financial support and for their belief in me;

My friends - Lauri, Ruth, Mary, Donna, Dinny, Andrea, JoAnn, Miguel, Ernie, Diana, Susana, Susan, Becky, Gail, Jessica, Bill, Elsa, Corinna Rose, Chris, Peter, Sandy, Rich, Judy, Steve, Glenda, Melinda and Ed for so much;

Jessica Murray for her beautiful art work;

Guillermo, Elsa, Carlin, Gilda, Maria, Pat, Chris, Juana, Larrie, Debbie and Lourdes for getting the data;

Barbara, Mark and David for teaching me how to massage the data;

Olivia, Aurora, Carol, Alicia, Nancy and all the teachers for access to their students.

Elinore Brown for her warmth as well as for those hours and hours and hours of typing;

Susan Hayes for always being there in a pinch.

Most importantly I want to thank the children in San José, California, El Paso, Texas, Walsenburg, Colorado, and Las Vegas, Nevada. This research is dedicated to them.

TABLE OF CONTENTS

		page
Precis	. . . . .	iii
ACKNOWLEDGMENTS	. . . . .	iv
LIST OF TABLES	. . . . .	ix
CHAPTER		
I	INTRODUCTION	1
	Background of the Problem	1
	Why Test Language Proficiency?	1
	Definition of the Research Problem	5
	Purpose of the Study	9
	Research Questions	11
II	CONCEPTUAL FRAMEWORK	12
	What Does It Mean to Know a Language?	12
	Description and Measurement of Bilingualism	17
	Bilingualism and Cognitive Development	24
III	METHODOLOGY	28
	Overview	28
	Subjects	29
	Sample Selection	32
	Instrumentation	33
	Measures of Linguistic Competence (oral/aural)	33
	Measure of Communicative Competence (oral/aural)	37
	Measures of Cognitive and Academic Functioning .	39
	Parent Questionnaire	42
	Data Collection Procedures	43
	Student Identification	43

		Page
	Test Administrators	44
	Field Testing	44
	Rater Training	44
	Scorer Training	46
	Samples #2 and #3	46
IV	RESULTS AND INTERPRETATION	50
	Bilingual Syntax Measure II: Reliability and Validity	50
	Inter-scorer Reliability	50
	Order and Tester Effects	51
	Internal Consistency	52
	BSM Levels vs. BSM Total Scores	53
	Communicative Competence Test: Reliability and Validity	58
	Inter-rater Reliabilities	58
	Order and Tester Effects	62
	Internal Consistency	63
	Question #1	66
	Question #2	70
	Question #3	75
	Equivalent Percentile Cutoff Procedure	76
	Grade-Level Regression Procedures	88
	Correlations	91
	Regression Analyses	95
	Factor Analyses	99
	Summary	103
	Question #4	104
V	CONCLUSIONS AND RECOMMENDATIONS	115
	Overview	115
	Conclusions	116

Page

Recommendations 121

BIBLIOGRAPHY 126

APPENDICES 143

Decision Rules for Sample #2 Selection 144
Distribution of Sample #2 . . . . . 145a
Sample #2 Crosstabulations of Spanish BSM II Level by English Level 145b
Teacher Communicative Competence Rating Scale . . . 146
Parent Permission Slip 147
BSM II Level Descriptions 149
Selected Examples of BSM II Items 150
CCT Puppets 151
CCT: Puppet Interaction Situations 152
CCT: Example of Tester Script 153
CCT: Example of Test Transcript 155
CCT: Story SequenceScene 156
CCT: Rating Scales 157
WNBD: Directions 164
WNBD: Pictures 166
CERAS: Examples from Administrator's Manual 168
CERAS: Examples from Student Response Booklet 170
CCS: Examples 171
Parent Questionnaire 172

LIST OF TABLES

Number		Page
1	LAU Language Proficiency Categories	2
2	Distribution of Sampled Districts and Classrooms . .	30
3	District Comparisons	31
4	CERAS Balance Categories of Structures Tested . . .	36
5	CTBS/S Level 1 - English	40
6	CTBS Level 1 - Spanish	41
7	MAT Primary II, Form G	42
8	Intraclass Correlation Coefficients for BSM II . . .	51
9	Analyses of Variance Showing BSM Tester and Order Effects	52
10	Split Half Reliabilities of BSM II	53
11	Intercorrelations of English BSM II With Measures of Language Proficiency and Achievement	55
12	Intercorrelations of Spanish BSM II With Measures of Language Proficiency	56
13	Correlations Between BSM Total Scores and Levels . .	56
14	Sample #1 Intraclass Correlations of CCT - 5 Raters .	58
15	Sample #1 Intraclass Correlations of CCT - 4 Raters .	58
16	Generalizability Coefficients for CCT	59
17	Score Distributions for BSM, CERAS, CCT and WNBD . .	61
18	Analyses of Variance Showing CCT Tester and Order Effects	63
19	Split-Half Reliabilities of CCT	63
20	CCT Subscale Correlation Coefficients Sample #1 - English	64 64 64

Number		Page
21	CCT Subscale Correlation Coefficients Sample #3 - English	64
22	CCT Subscale Correlation Coefficients Sample #1 - Spanish	65
23	CCT Subscale Correlation Coefficients Sample #3 - Spanish	65
24	Summary of Distribution of Semantically Equivalent Syntactic Items	69
25	Repeated Measures Analyses of Variance - Main Effects	72
26	Repeated Measures Analyses of Variance - Interaction Effects Language and Domain Variables	73
27	English Percentile Raw Score Equivlents and Limited vs. Proficient Cutoffs Using BSM II Levels as Criteria	78
28	Spanish Raw Score Equivalents and Limited vs. Proficient Cutoffs Using BSM Levels as Criteria . . .	79
29	Equivalent Percentile Procedure Crosstabulations of Limited/Proficient Level Designations Proficient = Level 6	80
30	Bilingual Classifications (RLP)	82
31	Results of Crosstabulations Between Bilingual Classifications Proficient=Level 6	83
32	Equivalent Percentile Procedure Crosstabulations of Limited/Proficient Level Designations Proficient = Level 5	84
33	Results of Crosstabulations Between Bilingual Classifications (RLP) Proficient = Level 5 . . .	86
34	Grade-Level Cutoff Scores	89
35	Grade-Level Regression Procedure Crosstabulations of Limited/Proficient Level Designations	90
36	Correlation Matrix of Language Proficiency Tests . .	93
37	Stepwise Multiple Regression of CTBS/S, Level 1 English-Total Battery Achievement on Proficiency Tests	95

Number		Page
38	Multiple Regression of CTBS English Total Battery on English Language Proficiency CCT Entered Last	97
39	Multiple Regression of CTBS English Reading and Language Totals on English Language Proficiency - CCT Entered Last	97
40	Multiple Regression of CTBS English Reading and Language Totals on English Language Proficiency-Linguistic Competence Tests Entered Last	98
41	English Rotated Factor Pattern: All English Language Tests, Teacher Ratings and CCS . . .	100
42	Communalities of Language Measures - English	101
43	Spanish Rotated Factor Pattern: Spanish Language Tests and CCS	101
44	Communalities of Language Measures - Spanish	102
45	Crosstabulations of Bilingual Classification by BSM II and CCT	105
46	Analysis of Variance Results Comparing 3 RLP Groups on the CCS, CTBS English Total Battery, Student Length of Residency in the U.S. and Parents' Educational Levels	106
47	Means for Three RLP Groups	107
48	Crosstabulations of Groups 1 and 3 on Student Place of Birth	108
49	T-Test Results of Comparisons of Mexican and U.S. Born Bilingual Students	109
50	Crosstabulations of Student Place of Birth With Student Language Use With Parents	110
51	Crosstabulations of Student Place of Birth With Parent Language Use With Student	110
52	T-test Comparisons of Language Proficiency and Cognitive Functioning Scores for Students Who Use Spanish Only (S) vs. Spanish and English (S & E) at Home	112

Number		Page
53	T-Test Comparisons of Student Language Proficiency and Cognitive Functioning Scores for Parents Who Use Spanish Only (S) vs. Spanish and English (S & E) at Home	113
54	T-test Comparisons of Student Language Proficiency and Cognitive Functioning Scores for Combined Student and Parent Home Language Use for Students Who Use Same Language(s) as Parents	114

LIST OF FIGURES

1	English CUTOFFS Using Level 5 and Level 6	87
2	Spanish CUTOFFS Using Level 5 and Level 6	87

CHAPTER I

INTRODUCTION

Background of the Problem

The use of language proficiency instruments to assess first (L1) and second (L2) language skills is of critical importance to the educational placement of non-English language background (NELB) students in the United States. These tests provide levels which indicate limited or proficient status in a given language. The levels are then used to decide upon the language of instruction, whether placement in English as a Second Language instruction is necessary, or if special language diagnosis and remediation are required. Once such decisions have been made, the levels may then be used as criteria for instructional grouping, placement, or determination of readiness to exit from special language programs. The numbers of students designated as "limited" in English (and/or L1) are also used to determine amounts of federal and state funding a school or district may receive.

Why Test Language Proficiency? Soon after the Lau v. Nichols Supreme Court decision of 1974, the Office of Civil Rights (OCR) prepared the LAU Remedies to help school districts comply with the decision. The LAU Remedies are recommended procedures for the sequential identification, assessment and educational placement of non-English language background students. Based upon the Office of Civil Rights' interpretation of the LAU decision, they request that school districts provide more equal

educational opportunity through use of the student's native language and English as a second language in schools. School districts were cited for non-compliance if their policies did not match these Remedies. During the first year of the Remedies, 90 percent of districts cited in non-compliance were so cited for failure to identify and assess English language proficiency of NELB students (OCR, 1978).

These Remedies depend heavily upon the use of language proficiency instruments to assess the L1 skills of NELB students. Specific LAU Language Proficiency Categories were developed which classify language proficiency into levels. Table 1 lists these categories.

Table 1

LAU Language Proficiency Categories

A.	Monolingual in a language other than English (non-English-Proficient (NEP))
B.	Predominant speaker of a language other than English, though he knows some English (limited-English-proficient (LEP))
C.	Bilingual, i.e., has equal facility in English and some other language (fluent-English-proficient (FEP) or limited-English-proficient (LEP))
D.	Predominant speaker of English, though he knows some other language (FEP)
E.	Monolingual in English, speaks no other language (FEP)

Most cases of non-compliance cited by OCR were due to the use of teacher judgment rather than standardized tests for determination of these levels. Because of the probable bias in teacher judgment, the LAU Remedies required school district personnel to use reliable and valid language proficiency instruments for determining LAU Categories and, consequently, the special instructional and language needs of NELB students.

During the five years following publication of the LAU Remedies, there appeared a proliferation of language proficiency instruments for use with NELB students. Many were created ad hoc, without proper field testing or validation; others were piloted on small numbers of students; only a few were thoroughly field tested prior to publication.

Along with the onslaught of language proficiency tests came a number of texts whose sole purpose was to evaluate the rapidly created instruments then deluging the market (DCBBE, 1975; Silverman et al., 1976; BABEL, 1972; Locks et al., 1978; Center for Bilingual Education, 1978; Pletcher et al., 1978). The conclusions of most of these evaluations indicated that there were few, if any, satisfactory tests available.

While school districts must comply with OCR recommendations for testing, the use of language proficiency instruments for measuring the level of English attainment of NELB students serves many other purposes. Where teachers desire diagnostic information, tests which provide detailed information about each child's attainment of particular English language skills may be used (Sánchez, 1976b). Such tests include those which accompany materials for teaching English as a Second Language (e.g., San Diego Continuum, Idea Kit) or those which can serve as diagnostic tools for use with a variety of language teaching materials (e.g., Mat-Sea-Cal).

Language proficiency testing is necessary not only for placement in special language programs and diagnosis of language needs, but also as an important component in decisions of when students are ready to function without special language assistance. The use of language tests for the purpose of exiting students from special language programs is referred to as <u>reclassification</u>.

To make reclassification decisions both state laws and federal recommendations suggest the use of multiple indicators of language proficiency. The purpose of reclassification is to determine when students are ready to function effectively in English-only curricula. Therefore, academic skills as well as oral English language proficiency should be assessed. According to the California Bilingual Education Improvement and Reform Act of 1980 (California Statutes, 1980), multiple criteria including achievement and oral language proficiency must be assessed before reclassification of NELB students proceeds. The Act implies that if a student is to function effectively in an English-only curriculum, s/he should be able to achieve in English as well as exhibit proficient oral language skills. Federal concern for the appropriate determination of academic and language skills necessary to function effectively in an all English-speaking classroom was indicated by the Request for Proposals Number 78-111: Development of Entry/Exit Criteria and Associated Assessment Procedures for Bilingual Education Projects (DHEW, 1978).

Oral proficiency tests do not necessarily indicate students' academic abilities in English. This is evidenced in correlations between oral language tests and standardized English achievement measures. Typically oral language is only moderately related to achievement. DeAvila, Cervantes and Duncan (1978) found simple correlations between oral language proficiency and CTBS Reading of .41 for grades 1, 3 and 5 and .49 for grades 7-12. Gillmore and Dickerson (1979) report Pearson correlations between five different measures of oral English proficiency and achievement which range from -.24 to .46. Ulibarri, Spencer and Rivas (1981) report correlations for grades 1, 3 and 5 between the three oral language tests and achievement. Correlations range from .08 to .47.

Only moderate correlations are expected if oral English proficiency is not sufficient to predict English academic achievement. DeAvila (1981) in a review and reanalysis of comparative language and achievement data reports that, of students classified as proficient in English "over half of the students . . . were also found to be below the 40th percentile in reading" (p. 13).

The distinction between oral language proficiency and language proficiency related to academic achievement is made at the theoretical level by Cummins (1981c, 1980) in his distinction between Basic Interpersonal Communication Skills (BICS) and Cognitive/Academic Language Proficiency (CALP). Arguing that language is not a unitary construct, he finds that most NELB children will acquire superficial BICS in the first two years of learning English. They will take longer, however, to reach grade level CALP achievement. Measurement of BICS may not provide any indication of the CALP skills in children. Since CALP skills are those necessary to succeed in schools, reclassification and exiting criteria should include CALP-type measures. In other words, tests which predict students' ability to function in English academically related skill areas are those which should be used to determine when NELB students can function in English-only instruction.

Definition of the Research Problem

While functional, achievement related language (CALP) tests might appear to be the most predictively valid measures, they have not, for the most part, been used in the assessment of NELB students. Rather, primarily linguistic ability (BICS) oriented tests, which measure surface features of language, have dominated (Dieterich, et al., 1979). On such tests,

total scores are converted to level designations and students, in turn, classified as limited or proficient in English.

These tests vary in content. The four most widely used instruments, the Language Assessment Scales (LAS) (DeAvila & Duncan, 1977), Bilingual Syntax Measure (BSM) (Burt, Dulay & Hernández-Chávez, 1975), Basic Inventory of Natural Language (BINL) (Herbert, 1977) and the Language Assessment Battery (LAB) (Board of Education, New York, 1976) were each designed to measure different aspects of language. The LAS samples items from receptive and productive phonology, lexicon, syntax and discourse (oral language); the BSM samples specific syntactic structures (oral); the BINL samples spontaneous speech (oral); the LAB samples various aspects of speaking, reading, writing and listening (oral and written). Although the tests measure different aspects and skills of language, and most measure only oral language skills, they are all used for the purposes of classifying NELB students into OCR acceptable LAU Categories and, subsequently, placement into specific educational programs.

If tests are used for the same purpose(s) they should, theoretically, agree in what they tell us about testees. In the case of language proficiency tests, they should agree in their level designations and student classifications. Comparative studies of these kinds of tests do not, however, indicate substantial agreement. Graham and Acosta (1979) found moderate correlations ranging from .39 to .82 between the BINL, LAB, LAS and BSM. They noted, however, that different tests produce different numbers of students classified as non-English speaking (NES), limited-English-speaking (LES), fluent-English speaking (FES), and proficient-English-speaking (PES). Agreement in classification between pairs of tests were 56% for the BINL and BSM, 65% for the BINL and LAS, and 77%

for the BSM and LAS. Gillmore and Dickerson (1979) compared classifications among five tests widely used in Texas. Included were the BINL, BSM and LAS as well as the Primary Acquisition of Language (PAL) and the Schutt Primary Language Indicator Test (SPLIT). Their results indicate lack of agreement among these tests to classify students as limited in English. Percents of agreement in classification as limited-English-speaking range from 45% for the BINL and BSM to 78% for the LAS and BSM.

Cervantes and Nakano (1979) compared language classifications on the BSM and LAS. Students in first and third grades were tested on both instruments. They found, on the average, that these two tests agreed on classification of NELB students only 14% of the time!

Finally, Ulibarri, Spencer and Rivas (1981) in their study for the California State Department of Education compared students' English level classifications on the BSM, LAS and BINL. Percents of agreement between test pairs ranged from 27% for the BSM and LAS to 65% for the BINL and LAS. Their findings suggest that tests will yield different numbers of students of limited English proficiency.

> . . . the NES/LES/FES categories cannot be considered equivalent across tests. In general, they did not agree in terms of the percentage of children classified as NES/LES/FES, and if they did, they did not agree in terms of which children were classified as NES/LES/FES.

Using the tests' raw scores, however, high correlations between tests were found. These range from .59 to .74. While correlations between raw scores might indicate that tests may be measuring the same abilities, the cutoff scores recommended for making classification decisions are not comparable from test to test.

What these studies have shown is that, according to classical test theory, the most widely used tests to measure English language proficiency of NELB students provide level designations which have a large error of measurement. Hence, we may not be correctly classifying NELB students. These tests may provide inaccurate indications of language skills and abilities. Consequently, we may make inappropriate placement decisions and not provide NELB students with equal educational opportunity, the intent behind both the LAU Decision and the OCR recommendations.

In 1980, the Department of Education and OCR proposed LAU Regulations (referred to as NPRM) which, though recently rescinded by Secretary of Education Terrell Bell, created even more of a need to use reliable and valid measures of language proficiency in making educational decisions about NELB students. The NPRM specified that procedures used to determine language proficiency should be:

> . . . generally regarded by experts on the assessment or analysis of language proficiency as capable of:
>
> (a) measuring the language skills necessary to make instructional or other educational decisions about these students; and
>
> (b) producing consistent results.

It also suggested that Relative Language Proficiency (RLP), or proficiencies in both L1 and L2, be measured before educational placement decisions are made. This was a shift from the English only testing requirements of the LAU Remedies. Most importantly, the NPRM required the use of multiple indicators of oral and academic language proficiency in making educational decisions about NELB students.

Although the revised LAU regulations were never adopted, and, hence, the LAU Remedies have remained in effect, a great deal of commentary

was produced around the proposed regulations. These included public testimony by thousands of school district personnel, language researchers and public officials who indicated the many problems enountered in testing and placement of NELB students. A LAU Testing and Assessment Implications Task Force was created by the Department of Education and the National Institute of Education to analyze these problems and make recommendations for the use of language proficiency tests in educational decision making. The document created by the Task Force, the Briefing Notebook (NIE, 1981), made specific recommendations for the improvement of language proficiency testing. Specifically, the use of multiple criteria for determining educational placement was recommended because of the predictive and concurrent validity problems of currently available language proficiency tests. "The single biggest problem with relying on one indicator of language proficiency lies in the technical limitations of all available procedures"(NIE, Executive Summary, 1981, p. 9).

One of the major conclusions reached by the Task Force was that the currently used instruments for measuring language proficiency of NELB students are inadequate.

> The various instruments are not telling us the same thing about students' language proficiency. With few exceptions, (they) do not identify the same percentage of children as NES/LES/FES nor are particular students reliably classified by these instruments (p. 75).

Purpose of the Study

Recent research on the assessment of bilingual language proficiency suggests that substantial numbers of bilingual students may be limited in both their native language and their second language (Skutnabb-Kangas &

Toukomaa, 1976; Cummins, 1979a; Jaakkola, 1977; Duncan & DeAvila, 1979; Burt & Dulay, 1980). Based upon results of various language proficiency measures, estimates in California report that as many as 94 percent of limited-English proficient Hispanic students are also limited in Spanish (Burt & Dulay, 1980). How much of this estimate is due to problems of instrumentation? Are such students truly limited in both L1 and L2? Or, are they misdiagnosed? Who are these limited proficient bilingual (LPB) students and what language skills do they possess? Are these students, who are often referred for special diagnosis and possible placement in special education programs, limited in cognitive and academic skills as well as linguistic skills? What might cause such limitations?

It may be that such students score low in both languages as a result of limited development in both languages or as a result of non-standard use. If bilingual students have been raised in diglossic environments where the first and second languages are used for functionally differentiated purposes (Ferguson, 1964; Fishman et al., 1971), measurement of language proficiency in either language may give only a partial view of their true language proficiency. Perhaps their total verbal repertoire is distributed across two languages with the result that scores in either of the languages are considerably below monolingual norms. Or, perhaps their language proficiency is inadequately measured by linguistic ability tests which do not give a true picture of their functional language skills.

It is the purpose of this study to examine the language proficiencies of limited proficient bilingual, Mexican-American, third-grade children in San José, California, the area from which Burt and Dulay (1980) report their highest incidence of LPB students. Specifically, this study will

investigate whether:

(a) verbal skills of limited proficient bilingual students are distributed across languages in such a way that measurement in only one language gives the illusion of a verbal deficit;

(b) students are classified as limited proficient bilingual on functional (communicative) as well as linguistic measures.

Research Questions

1. Among students classified as "limited proficient bilingual," is the total repertoire of grammatical skills distributed across languages? If so, is this distribution complementary?

2. Among students classified as "limited proficient bilingual," is the total repertoire of vocabulary and communication skills distributed across languages as a function of domain?

3. Among students classified as "limited proficient bilingual," is there a difference in assigned bilingual proficiencies when different types of proficiency tests (linguistic competence vs. communicative competence) are used to assess verbal behavior?

4. What variables other than language proficiency (i.e., cognitive functioning, achievement, length of residency) differentiate between limited proficient bilingual and non-limited proficient bilingual students?

CHAPTER II

CONCEPTUAL FRAMEWORK

A description of the language characteristics of "limited proficient bilingual"* students is dependent upon precise definition and theoretically valid measurement of bilingual language proficiency. Unfortunately, there is a lack of agreement concerning what language skills bilinguals possess as well as how such skills can be measured. A "limited proficient bilingual" may be generally defined as one who lacks native-like proficiency both in the first and second languages. This imprecision reflects the general problems in the definition of "what it means to know a language" compounded by those associated with the description and measurement of bilingualism (Kelly, 1967; Macnamara, 1967; DeAvila & Duncan, 1978a).

What Does It Mean to Know a Language?

Language proficiency tests are most often based on a behavioral-structural model (Valette, 1967, 1977; Lado, 1961; Harris, 1969), which views language as "a linguistic phenomenon" which "must be broken down into its linguistic components" in order to be tested (Davies, 1978). Therefore, language proficiency has been divided into skill areas: speaking, listening, reading and writing. Within each skill are specific, discrete elements of language: phonology, lexicon, syntax and semantics. Tests which measure language proficiency based upon this model of language are

*Limited proficient bilingual students are often referred to as comparably limited or limited English dominant bilinguals.

referred to as tests of linguistic competence.

In order for these measures to serve as valid indicators of overall language proficiency, all skill areas and elements should be sampled. In general, however, practical constraints (i.e., time, age of student) prohibit such elaborate testing. Rather, tests sample items from one or more areas and scores are assumed to indicate overall "language proficiency." This practice is justified on the grounds that there exists a unitary construct called "language proficiency" (Oller & Perkins, 1980). Citing the interrelatedness of language subskills and the general tendency of these subskills to load on a "global language factor," Oller provides evidence for this single underlying factor theory (Oller & Perkins, 1978; Oller & Perkins, 1980).

In recent years (late 1960s to present) the measurement of language proficiency which reflects only linguistic competence has been criticized for sampling "decontextualized" language (Carroll, 1971; Jakobovits, 1970; Shuy, 1978; Rosansky, 1980; Schulz & Bartz, 1975; Savignon, 1972; Briere, 1971; Erickson & Omark, 1981; Sánchez, 1976b). Carroll (1971) first advocated the use of tests which would measure "levels of performance which are relevant to the kinds of situations in which the examinees will find themselves" (p. 319). Because language ability is assessed out of the context of real-life communications, only approximate indications of proficiency are provided. Measures of linguistic competence are unable to capture language as it is actually used in communication.

> Language is not a set of unrelated bits . . . it forms a whole . . . the bits must be integrated and tested in combination with one another . . . language learning is purposeful . . . the purpose is always communicative . . . what must be tested is communicative ability not formal knowledge. (Davies, 1978:215)

A new view of language proficiency which focuses on language used in social settings has emerged. Behaviors which reveal that an individual can communicate effectively are now included in theories (Canale & Swain, 1980) and definitions (Wiemann & Backlund, 1980) of communicative competence. Within this emerging field can be seen the development of tests which attempt to measure the functional use of language (Jones & Spolsky, 1975; Savignon, 1972; Schulz & Bartz, 1975; Overall, 1978; Johnson, 1980; Legaretta, 1979; Politzer et al., 1981; Cummins, 1981b).

Communicative competence expands the limited view of language as linguistic competence to include knowledge of when to use what forms of language to whom.

> The members of a speech community share not only linguistic competence, the ability to understand and produce the theoretically infinite set of sentences comprising the language, but also communicative competence, the knowledge of when to speak and when to remain silent, and what to say to whom and when. (Hymes, 1972)

Because measures of communicative competence go beyond linguistic features of language and attempt to document actual functional language use, they are believed to be more valid indicators of language proficiency than tests of linguistic competence (Jacobovits, 1970; Davies, 1978; Briere, 1971; Rosansky, 1980; Morrow, 1977; Sánchez, 1976b).

Linguistic competence and communicative competence share some common properties; many speakers who are communicatively competent may also be linguistically competent. However, while these constructs are obviously related, they are not direct indicators of each other. Clark (1972) comments on the "absence of a close and easily determined relationship between sheer linguistic ability . . . and communicative proficiency. In many studies, positive correlations have been found between measures

of linguistic competence and communicative competence. Carroll (1967) correlated the Foreign Service Interviews (FSI) with the MLA proficiency tests and found a range of correlations between .58 and .86. Clark (1972) correlated the FSI with measures of vocabulary and grammar. He found correlations between .82 and .92. Schulz and Bartz (1975) found a wide range of correlations, .31 to .74.

On the other hand, however, some studies report that competence in grammar (linguistic competence) does not reliably predict competence in communication (communicative competence). Tucker (1974) found that students who varied in linguistic competence were indistinguishable in their communication skills. Ervin-Tripp's (1967) research led her to the conclusion that "second language learners can successfully communicate in spite of quite deviant grammars." Overall (1978) in her study of 4th and 6th grade Mexican-American children provided further evidence that a range of relations exists between linguistic competence and communicative competence; her findings indicate that these relations could best be described as curvilinear. Specifically, after examination of scatterplots of scores on both measures, she found that students with low levels of linguistic competence varied in their degree of communicative competence.

In a recent study conducted by Politzer, Shohamy and McGroarty (1981), Mexican-American high school and elementary students were administered linguistic and communicative tests. A range of correlations, between .16 and .67, were found. The researchers caution interpretation of linguistic competence and communicative competence as similar constructs. Rather, through examination of scatterplots, they found that "a certain amount of linguistic competence is the prerequisite for communicative competence

and even at relatively low or relatively high levels of linguistic competence, there can be considerable variance in communicative competence" (Politzer, et al., 1981).

Palmer (1979), Farhady (1979) and Cziko (1981) cite differences between tests of grammar and tests of communication for different types of learners. For some students, perhaps those who have learned the second language in formal settings, there exists a high, linear correlation. For others, perhaps those who have learned their L2 in informal, non-classroom settings, the relationship is less consistent and direct.

For the placement of NELB students in the United States, language assessment instruments of linguistic competence dominate (Silverman et al., 1976; Pletcher, 1978; Center for Bilingual Education, 1978; DeAvila & Duncan, 1978a; Dieterich, et al., 1979). These tests usually measure, with varying degrees of success, a small, limited portion of the students' total verbal repertoire. The tests are, however, easy to administer, relatively easy to score and cost effective. Apparently, tests are chosen because they suit "administrative needs rather than pedagogical ones" (Shuy, 1978).

Complicating this difference between linguistic and communicative competence tests is Cummins' (1981c, 1980) distinction between Cognitive Academic Language Proficiency (CALP) and Basic Interpersonal Communication Skills (BICS). Most linguistic and communicative tests do not appear to measure achievement related skills (CALP). Rather, they may only measure linguistic and communicative surface features of language (BICS).

Due to the relatively recent development of the notions of communicative competence, BICS and CALP, the use of measures of achievement related communication skills as indicators of language proficiency has had a slow and barely visible beginning. The important decisions regarding placement

of NELB students are made with results of primarily linguistic, BICS type measures. For example, none of the California state adopted tests of language proficiency for use in placement of NELB students measures communicative competence. Moreover, none of these tests appears to sample school related language tasks.

Finally, while both linguistic and communicative competence measures may provide different and important information about students' language proficiencies, the use of one or the other for placement of NELB students should be governed by the answer to the following question, appropriately posed by Politzer et al. (1981): Which aspects of language proficiency—linguistic competence or communicative competence—can best predict success in school?

Description and Measurement Of Bilingualism

The measurement of language proficiency is a complex endeavor. The difficulties associated with design and development of monolingual measures of linguistic and communicative competence are magnified when the measurement of bilingual language proficiency is attempted.

> A needs analysis conducted by the National Institute of Education (Locks et al., 1978) showed that there continued to be a persistent need for tests assessing the language competence of limited English speaking pupils. A glance at the proceedings of the latest International Conference on Frontiers in Language Proficiency and Dominance Testing reveals that existing tests measuring relative dominance of (Spanish/English) bilinguals tend to be problematical and/or unsatisfactory (Morales, 1978) and that almost all the important theoretical issues concerning the language testing of bilinguals and the establishment of their language dominance configurations remain unsettled. (Politzer, 1979)

The problems that plague developers and users of bilingual language proficiency measures are many. Basic to these problems is disagreement as to what constitutes language proficiency and, therefore, bilingual language proficiency. Parker, in her attempt to define bilingualism, poignantly sums up this problem: "Bilingualism indicates some degree of proficiency in a second language. Researchers are discovering the complex nature of what constitutes language 'proficiency.' Arbitrary definitions of language 'proficiency' are adopted for particular purposes such as eligibility for a bilingual program" (Parker, 1978, p. 5).

Definitions of bilingualism range from possession of at least one language skill even to a minimal degree in the second language (MacNamara, 1967) or the ability to "produce complete, meaningful utterances in the other language" (Haugen, 1953), to the "alternate use of two languages" (Weinreich, 1953), to "native-like control of two languages" (Bloomfield, 1933) and "active, completely equal mastery of two or more languages" (Braun, 1937, translated by Haugen (1968) cited by Toukomaa & Skutnabb-Kangas, 1977).

Bilinguals have recently been characterized as speakers "who know both languages equally well" (Hernández-Chávez et al., 1978) or as "any individual who lives in a home where a first language other than English is spoken by one or more members of the household" (Burt & Dulay, 1978).

While the "ideal" balanced bilingual might possess equal skill in all aspects of two languages, it is generally "intended in a more limited sense" as a person who is balanced in "some particular facet of linguistic performance" (MacNamara, 1967). For example, a grammatically balanced bilingual whose second language knowledge of vocabulary is confined to

what s/he learns in school may fail to learn many words which native speakers of the language know prior to entrance in school. While s/he may be balanced grammatically, s/he may not be balanced in vocabulary. Thus, variations in an individual's degree of bilingualism can occur within the same skill area (i.e., speaking) but across elements (grammar vs. vocabulary). It can also vary across skills (Weinreich, 1953).

Complicated by the issue of definition are the questions of what to measure and how to measure it. Theoretical investigations of bilingualism reveal that the use of two languages involves a complex set of skills which are difficult to measure (Kelly, 1969; Mackey, 1962; MacNamara, 1967; DeAvila & Duncan, 1978a). Consequently, direct measurement has often been replaced by semidirect and indirect measures (see MacNamara, 1967; Clark, 1975).

Direct measurement of language skills would entail documentation, in real-life-like situations, of the use of language. The difficulty of thus capturing language proficiency in non-artificial situations has led language researchers to rely upon indirect indicators of proficiency. Tests of fluency (speed), flexibility and word preference as well as self-report instruments have been used as indirect indicators of linguistic or communicative competence. Although such measures have been found to be highly interrelated (Lambert, Havelka & Crosby, 1958) their use as valid indicators of bilingual language proficiency has been seriously questioned (MacNamara, 1967; Kelly, 1969; Hernández-Chávez, 1978). The fact that indirect measures of language proficiency correlate may indicate that they measure the same thing. However, "it remains to be demonstrated that . . . whatever factor it is that is being measured here bears a close relation to

more direct measures" (Hernández-Chávez, 1978).

While tests may directly or indirectly measure monolingual and bilingual language proficiency, these measures can be further differentiated, and their comparability and validity complicated, by their reliance on discrete point or integrative items (Jones, 1975; Cziko, 1981). Discrete point tests are composed of items which deal with individual language elements (grammar, vocabulary, phonology). Each point is scored separately. Linguistic competence tests are typically of this type. Such discrete point tests, although they may be reliable and possess construct validity, must be validated against appropriate criteria "because the items of a discrete point test typically involve tasks that bear little resemblance to tasks of actual language use" (Cziko, 1981:36).

Integrative tests of language proficiency are those which integrate many points into the scoring of items. Cloze, dictation and the interview are types of integrative tests. These tests appear to require use of language which is necessary in real-life, meaningful situations (Oller, 1979). Perhaps, as Cziko (1981) points out, it is the underlying cognitive process which makes performance on integrative tests similar to real life language use. In both situations the student must use knowledge of "the syntactic, semantic, discourse, sociolinguistic and extralinguistic rules of language behavior to formulate what he is about to say or write as well as to predict what he is about to hear or read" (p. 38). Integrative tests are more direct measures of language proficiency.

In general, linguistic competence tests tend to fall into the discrete point, indirect categories. Tests of communicative competence, on the other hand, have been developed as more integrative, direct measures.

Apart from the above mentioned validity issues relating to direct, indirect and discrete point or integrative tests, there are special problems involved in testing for equivalent proficiencies in two languages. It is not like the double of proficiency testing in a single language (Hernández-Chávez, 1978). Equivalent tests in two languages, while statistically parallel due to field testing, may not be truly parallel. If one language test form is a translation of the other, lack of parity between features of the two languages make the forms suspect. Davies (1978) citing Valette (1977) explains this inequivalence between languages: "structural linguistics was used as a basis for contrastive analysis in order to identify the problems faced by the second language learner; these problems . . . are patterns and units that have no counterpart in the native language or a different distributional pattern" (Davies, 1978:148).

Built into the use of translations is the fact that scoring systems, while parallel across both languages, may not be equivalent, due to similar inequivalencies of languages. For example, in some languages different vocabulary and grammatical structures vary in their frequency of occurrence. To use translations of vocabulary or grammar items and then to score the "equivalent" forms as if the item held the same weight in each language may be erroneous.

Perhaps the most overlooked yet pervasive problem associated with the measurement and description of bilingual language proficiency is the accurate assessment of language used by students raised in bilingual communities in which the first and second language may be used for societally differentiated functions. This separation of language by functions, at the societal level, has been labeled <u>diglossia</u> (Ferguson,

1964). For example, in many bilingual communities in the United States, the home language (low variety) is used for community, informal purposes; the second language (high variety) is used for official business (i.e., work, business, school, etc.).

> . . . all speech communities have complex repertoires of language behavior. Understanding this variation and complexity is crucial for the analysis of bilingual education. Indeed, bilingual education programs arise in response to problems of language variation and repertoire. (Ferguson, 1978)

Because of the existence of more than two varieties in both English and Spanish, Peñalosa (1980) suggests the use of the term polyglossia to describe Chicano communities.

Most research on bilingual language proficiency does not take into account the bilingual's varying use of languages (Cohen, 1975). Some investigators (Fishman, Cooper & Ma, 1971; Cohen, 1975; Politzer & McKay, 1975; Zirkel, 1976; Poplack, 1981; Rodríquez-Brown & Olivares, 1981) have attempted to define and measure language proficiency as a function of context. Using the domains of home, neighborhood, religion and work (for adults), these studies revealed that Spanish/English bilinguals, in many communities in the United States, are more proficient in Spanish in the traditional domains of home, neighborhood and religion and more proficient in English in work and education. The need for language proficiency measures which take into account the notion of context (domains) is supported when relative language proficiency of bilinguals is examined in relation to domains. If language proficiency

> . . . had not been broken down by domain, performance would have been described as balanced in terms of differences between English and Spanish average scores. Yet . . . these groups exhibited significant differences between English and Spanish average scores in one or more domains. (Cohen, 1975)

Thus, for bilingual children in many United States Hispanic communities, language proficiency reflects a functional distribution of two languages.

> The notion of balance (or of dominance viewed as lack of balance) would seem to be even less realistic, precisely because socially patterned bilingualism can exist as a stabilized phenomenon only if there is functional differentiation between two languages rather than merely global dominance or balance. From the point of view of sociolinguistics any society that produces functionally balanced bilinguals (i.e., bilinguals who use both their languages equally and equally well in all contexts) must soon cease to be bilingual because no society needs two languages for one and the same set of functions. (Fishman,et al., 1971:560)

If bilingual students differentiate their use of the languages, that is, they use their first and second languages in different situations, the accurate assessment of their language proficiency must take situation into account.

Thus, many bilingual students use language(s) differently from their monolingual peers. If this is the case, tests normed on monolingual students may be inappropriate "because the language familiar to non-English speaking children is often a combination of two languages" (DeAvila & Havassy, 1978; see also Gumperz & Hernández-Chávez, 1970; Hernández-Chávez, Cohen & Beltramo, 1975; Elías-Olivares & Fallis, 1979; Barker, 1974; Zentella, 1979; Peñalosa, 1980, for descriptions of language mixing in bilingual communities).

> Specification of the norm, or the problem of deciding which utterances are grammatical and acceptable as a part of the basic core of language common to the community as a whole, is always a complex problem. It is doubly complex, however, when the community of speakers is a bilingual one, for we must speak of a bilingual code in addition to the two separate codes available to the speaker, whose alternative use by this speaker forms the environment for the creation of the third bilingual code. (Albin, 1972:27, emphasis added)

If bilingual language proficiency is judged according to comparisons with monolingual standards, bilinguals may appear limited because they are judged on just part of their bilingual repertoire (Lavandera, 1979). When bilingual students appear as limited in gammar in both languages, for example, "This appearance may actually be due to the test developer's lack of knowledge of the rules of the grammar used by the students" (Labov, 1970).

The proposed research is an initial attempt to describe bilingual language proficiency of Mexican-American children who may have been raised in diglossic bilingual communities and who are found to be "limited" in Spanish and English language skills. Because results of proficiency assessment are used to place bilingual children in appropriate classroom situations (i.e., bilingual or monolingual English) the importance of correctly identifying children is magnified. When "language" is used as a basis for comparison, predictions of achievment potential, grouping for instruction, eligibility for funding, etc., careful examination of language and its measurement are major concerns (Christian & Tripp, 1978).

Bilingualism and Cognitive Development

Along with reports of "limited bilingual proficiency" or students with "dual language limitations" goes the assumption that lack of proficiency in both L1 and L2 indicates cognitive limitations as well. Anecdotal evidence is available to support this assumption; the incidence of limited proficient bilinguals (sometimes referred to as "alinguals") in Special Education classes is common.

> The extremists here come to believe that Chicanos are alingual, that is, that they have no native language, no language at all. In effect, this type of instructor questions the intelligence of the Chicano, his power of conceptualization and his ability to reason logically. (Sánchez, 1976a)

Throughout the history of research on bilingualism, we find conflicting reports of both the positive and negative effects of experiences with two languages. Early studies (Saer, 1923; Darcy, 1946) tended to find negative cognitive consequences of bilingualism. Later studies (Peal, 1961; Peal & Lambert, 1962; Anisfeld, 1964; Torrance et al., 1970; Balkan, 1970; Ben-Zeev, 1972; Ianco-Worrall, 1972; Cummins & Gulutsan, 1973; Merino, 1975; Ramírez et al., 1977 and Duncan & DeAvila, 1979) have found positive effects. While the early studies have been criticized for lack of methodological rigor (i.e., socioeconomic status or language proficiency not controlled), later studies, though improved, were not faultless (Paulston, 1979a).

The general findings at the present time indicate that proficiency in two languages can have positive effects on cognitive development. Anisfeld (1964) theorizes that bilingualism may influence thinking in several ways; it may give the individual slightly different views of the world, a more flexible approach to problem solving and may encourage him to think abstractly at an earlier age than would a monolingual.

Results of research from around the world (Peal, 1961; Peal & Lambert, 1962; Anisfeld, 1964 from Canada; Torrance et al., 1970 from Singapore; Balkan, 1970 from Switzerland; Ianco-Worrall, 1972 from South Africa; Ben-Zeev, 1972 from Israel and New York; Cummins & Gulutsan, 1973 from Western Canada, and Duncan & DeAvila, 1979 from the United States)

indicate that proficient bilinguals have "a more diversified structure of intelligence . . . more flexibility in thought" (see Lambert, 1977 for review). Bilingual students show definite advantages over monolingual students in terms of cognitive flexibility, creativity and divergent thought.

Cummins (1979a) has tried to explain the differences between the early "negative" and later "positive" findings. He attributes the differences to threshold levels of language proficiency which bilinguals must attain before positive consequences of bilingualism will be found. This threshold hypothesis has been discussed by Toukomaa and Skutnabb-Kangas (1977) in describing language proficiency of Finnish minorities in Sweden.

> The basis for the possible attainment of the threshold level of L2 competence seems to be the level attained in the mother tongue. If in an early stage of its development a minority child finds itself in a foreign-language learning environment without contemporaneously receiving the requisite support in its mother tongue, the development of its skill in the mother tongue will slow down or even cease, leaving the child without a basis for learning the second language well enough to attain the threshold in it (1977:28).

If L1 skills reach the abstract level, they claim, students will be better able to abstract and master L2 concepts.

Cummins has explained contradictory findings of positive and negative effects of bilingualism by citing the threshold hypothesis. The negative results, to a great extent, were found among language minority students whose L1 was gradually being replaced by L2 ("subtractive" bilingualism). In these situations, children developed low levels of language proficiency in both languages. In most of the studies where cognitive advantages have been found, students continued to develop their L1 while acquiring L2 ("additive" bilingualism).

These data have led to the hypothesis that there may be threshold levels of linguistic proficiency which bilingual children must attain both in order to avoid cognitive deficits and to allow the potentially beneficial aspects of becoming bilingual to influence cogntive growth. The threshold hypothesis assumes that those aspects of bilingualism which might positively influence cognitive growth are unlikely to come into effect until the child has attained a certain minimum or threshold level of proficiency in the second language. Similarly, if bilingual children attain only a very low level of proficiency in one or both of their languages, their interaction with the environment through these languages both in terms of input and output is likely to be impoverished.

Support for this hypothesis to expalin the effects of bilingualism on cognitive development among Mexican-American students in the United States comes from studies by Duncan and DeAvila (1979). In their comparison of monolingual and bilingual students they found that language minority students who were proficient in both L1 and L2 performed significantly better on cognitive tasks than either monolinguals, partial bilinguals (students who have not yet become proficient in L2), and those bilinguals who were "limited" in proficiencies in both L1 and L2.

CHAPTER III

METHODOLOGY

Overview

The purpose of this investigation is to describe relationships among types of language proficiency measures and language proficiency classifications in Spanish/English bilingual students. Secondary analyses comparing these variables, measures of cognitive and academic functioning and home background variables are also performed. Of specific interest are two groups of bilingual students: (1) limited proficient bilinguals (LPB), those who appear to be limited in both L1 and L2; and (2) proficient bilinguals (PB), those who appear to be proficient in both L1 and L2. While LPB students are the target population, a control group of PB students was selected for comparisons with demographically similar, though linguistically different, students.

Non-monolingual Mexican-American students from 25 third-grade classrooms in four Santa Clara County school districts were the selected population. Third graders were selected because of the need to make placement and reclassification decisions at this level. Three samples were chosen.

Sample #1 was composed of 21 students judged high in communication skills in one language and low in the other; an author designed test of communication skills was field tested on this sample; they also were given the Bilingual Syntax Measure II. Recordings of both tests, in English and Spanish, were rated or scored.

Sample #2 was composed of 173 students whose previously collected dual language proficiency data indicated that they might be limited proficient or proficient bilinguals; this sample was administered the Bilingual Syntax Measure II in both English and Spanish.

Sample #3 was a subsample of Sample #2 which included students whose linguistic competence skills (BSM II scores) indicated limited bilingual proficiency (N=43) or bilingual proficiency (N=38). These students took a battery of linguistic, communicative and cognitive tests. Their linguistic tests were scored, the communicative test was rated. Demographic data were also collected on this sample.

All students were given a final linguistic competence test (the CERAS Balance Grammar Subtest) in both English and Spanish. After the close of school, district-administered achievement test scores were collected on students in the three samples.

Subjects

Mexican-American third grade students in Santa Clara County School districts (San José, California) were the selected population for this investigation. These districts have the highest reported percent of LPB students in California. Dulay and Burt (1980) report that 94% of the limited English proficienct students from bilingual program classes in Santa Clara County were found to be limited in Spanish as well. Although critics of this research cite faulty methodology and reporting procedures (DeAvila & Ulibarri, 1981) the reported preponderance of LPB students warrants investigation. Hence, this research sampled students from Santa Clara County to study and describe the nature of limited bilingual language proficiency.

Two large and two small school districts participated in the study. Twenty-five third grade classrooms in 18 schools were selected. The distribution of classrooms by schools and districts can be found in Table 2. Classes contained at least five and as many as twenty-five non-monolingual Mexican-American students.

Table 2

Distribution of Sampled Districts and Classrooms

District	# of Schools	# of Classrooms
#1	5	8
#2	2	2
#3	8	12
#4	3	3
	N = 18 schools	25 class-rooms

There is slight variation among districts in SES, percent minority, percent Spanish language background, funding sources and types of bilingual instructional designs. Table 3 displays these characteristics. Most of the students (87%) were sampled from two districts, one inner city and the other suburban, with quite varied bilingual instructional programming and funding sources. Classes varied from those with bilingual teachers and bilingual instruction (both Title VII and non-Title VII funded) to those with monolingual teachers in English only programs. A small number (13%) of students were sampled from small suburban districts with similar demography and bilingual program design.

Table 3

District Comparisons

District	SES	Percent[1] minority	Percent[4] Spanish	Title I[2]	Title VII[1]	Percent[3] Bilingual	N of Students	% of Total
#1	inner-city lower class blue collar	81	60.5	yes	80	100	86	42.5
#2	small suburban lower/middle class migrant farm labor and professional	50+	8.0	no	100	100	9	4.0
#3	large suburban lower class blue collar farm labor	85 (7 schools) 55 (1 school)	44.0	yes	none	58	90	44.5
#4	small suburban lower/middle class blue collar	56	36.0	no, but eligible	100	100	18	9.0
						Total N of Students =	203	(100%)

[1] average percent from sampled schools

[2] Title I eligibility is usually determined by a combination of factors: percent of Aid to Families with Dependent Children (AFDC); average school achievement below the 50th percentile; percent of limited English speakers.

[3] average number of sampled classrooms with some form of bilingual instruction.

[4] average percent of Spanish language background students enrolled in sampled schools.

The average age of students sampled was eight years old with a range from seven to eleven years. The large spread is due to several factors: (1) multi-graded classrooms contributed 7, 10 and 11 year-olds who were working at the third grade level; (2) older students may have been retained or were enrolled in lower grades than age required upon entering the United States.

Students were not differentiated by sex. Results of research with similar students (Ulibarri, Spencer & Rivas, 1981) indicate no differences between sexes on language proficiency tests.

For a subsample of students (Sample #3, N = 81), parents' educational levels were obtained. The average number of years that mothers went to school was 5.8 with a range from 0 to 13 years and standard deviation of 3 years. Fathers went to school for 5.9 years, on the average, with a range from 0 to 15 years and standard deviation of 3.46.

Sample Selection

Three samples of students were chosen: Sample #1 (N=21) was selected on the basis of teacher judgments of student communicative competence skills and dual language proficiency data. Neither PB nor LPB students were selected for Sample #1; rather, in order to maximize the number of LPB and PB students available for Sample #2 and Sample #3, fifteen students judged as high in Spanish/low in English and six students judged as high in English/low in Spanish were chosen. The Communicative Competence Test (CCT) was field-tested with this sample. Sample #2 (N=173) was composed of a group of students, pre-selected on the basis of dual language proficiency data already available in the schools, half of whom appeared to be limited proficient bilingual (LPB) and half apparently proficient

bilingual (PB)*. Sample #2 was administered a linguistic competence test in Spanish and English. Sample #3 (N=81) was a subsample of Sample #2, chosen on the basis of English and Spanish scores received on the linguistic competence test administered to Sample #2. Analyses to answer the research questions were performed on Sample #3 data.

Instrumentation

Four types of instrumentation were used in this study: (1) linguistic competence measures; (2) a communicative competence test; (3) cognitive and academic functioning measures, and (4) demographic indicators. Linguistic competence skills were measured by the Bilingual Syntax Measure II (Burt, Dulay, Hernández-Chávez & Taleporos, 1980), the CERAS Balance Revised Grammar Subtest (Ramírez & Politzer, 1975, revised by Hayes, 1981) and a Word Naming By Domain task. Communicative competence skills were measured by an author-designed instrument that was field tested on Sample #1. Cognitive functioning was assessed with the Cartoon Conservation Scales (DeAvila, 1980a); academic achievement was measured through district administered norm referenced tests. A parent questionnaire was designed and used to obtain specific demographic data (see Appendix for samples of instrumentation).

Measures of Linguistic Competence (oral/aural). The Bilingual Syntax Measure II (Burt, Dulay, Hernández-Chávez & Taleporos, 1980) is an individually administered instrument used to identify a student's control of basic grammatical structures in English and Spanish. There are 22 items in English and 21 in Spanish. Students are asked questions that refer to

*See Appendix for selection criteria.

one or more pictures. Answers are scored for grammaticality, not for content. Scores are totaled and categorized according to the levels into which they fall. There are six levels; Level 6 is the highest and denotes fluent language proficiency. The authors have determined that limited proficiency students are those who fall at Levels 1 through 5, or at Levels 1 through 4, depending upon district English norms. The BSM II, the upper level form of the original BSM, is geared for students in grades 3-12.

Test-retest reliability for grades 4-8 (N=85) is expressed as correlations of subcluster scores and level assignments. Correlation coefficients (type unspecified) were approximately .83 for English and .93 for Spanish. Internal consistency is reported for subcluster and total scores. Average English Alpha = .85 and Spanish Alpha = .75. Concurrent and predictive validities are not reported (Burt, Dulay et al., 1980). Inter-scorer reliabilities are not reported. Analysis of Sample #2 BSM data provided additional technical information. This is reported in Chapter Four.

Although there has been criticism of the reliability, content and predictive validity of the test (Cervantes & Nakano, 1979; Ulibarri, Spencer & Rivas, 1981; Rosansky, 1979; Sánchez, 1976b) it was chosen for this study due to its widespread popularity as an assessment instrument for classification of NELB students; it is usually used to determine eligibility for entrance into monolingual or special language programs.

Center for Educational Research at Stanford: Bilingual Balance Test, Grammar Subtest. Developed by Ramírez and Politzer (1975), the CERAS Balance measures student control of ten grammatical structures. The English grammar test contains structures which "are known to contain some production problems for speakers of Spanish. The Spanish version was a

translation of the English version. In other words, the Spanish version was not based on particular linguistic items but provided a general way of comparing proficiency in grammar in two languages" (Politzer & McKay, 1975).

Items may be elicited in individual interviews or in group settings. Students are asked questions which refer to one or more pictures. In the group administration (used in this investigation) students are provided with a one sentence descriptor of one picture and then asked for the correct descriptor of an adjacent picture. "This girl is singing. What are they doing?" (pointing to girls singing). Then, three possible answers, one of which is correct, are read. The revised version of this subtest used modifications made by California State University at Los Angeles, Title VII Evaluation Dissemination and Assessment Center in conjunction with a format created by this researcher. This final edition was designed for ease in group administration. Each page of the test contains an example of the item on the top (i.e., conversion of singular to plural); the example is read and the answers reviewed with students. The succeeding two items, which measure the same type of conversion, are then given. When the tester reads the answer choices, students circle the appropriate corresponding letter. Answer choices are read twice.

Reliability of the unrevised Spanish Grammar test is reported as Cronbach's Alpha = .94; the English grammar Alpha = .96. Concurrent and predictive validity data are not available. Content validity indicates that this instrument appears to be a receptive test of control over ten syntactic structures (listed in Table 4); each structure is measured with two items.

Table 4

CERAS Balance Categories
of Structures Tested

I.	singular to plural
II.	plural to singular
III.	present tense to past tense
IV.	past affirmative to present negative
V.	location—prepositions of place
VI.	questions, indirect to direct
VII.	commands, indirect to direct
VIII.	questions, direct to indirect
IX.	commands, direct to indirect
X.	comparatives

Word Naming by Domain (WNBD). This productive vocabulary testing technique was developed by Fishman, Cooper, Ma et al. (1971) for a study of language use in a Puerto Rican community in New Jersey. It was used with Mexican-American communities by Cohen (1975). The technique involves asking students to name objects commonly found in settings associated with the domains of home (kitchen), education (school), religion (church) and neighborhood (street). The student is given 45 seconds to name as many words as s/he can for objects found in specific settings. Testing is first in one language and then in another.

The technique was adapted for use in this study: (1) the time limit was extended to two minutes; (2) pictures of home and school settings were used to elicit vocabulary; (3) only the domains of home and school were measured; (4) examiner prompts ("Anything else you can think of?") were allowed, and (5) vocabulary was noted rather than simply counted with

hashmarks. The nature of this technique does not lend itself to conventional reliability analysis. It appears to be a valid indicator of students' productive control of domain-specific vocabulary. (Copies of stimulus pictures are found in Appendix.)

Measure of Communicative Competence (oral/aural). Communicative Competence Test (CCT). In order to answer the research questions, it was necessary to measure communications skills. A test measuring these skills was developed by the researcher. Communicative competence tasks were chosen from those used in previous research measuring communication skills in children (Overall, 1978; Johnson, 1980; Legarreta, 1979; Grimm, 1975; Politzer, Shohamy & McGroarty, 1981; Flavell et al., 1968; Krauss & Glucksberg, 1969; Bennett & Slaughter, 1981).

The development of a working definition of "communicative competence" and, therefore, communication skills which, when measured would indicate such competence, proved quite difficult. Although "communicative competence" is defined (Wiemann & Backlund, 1980), and theories of communicative competence" and its dimensions proposed (Morrow, 1977; Canale & Swain, 1980; Canale, 1981; Allen & Brown, 1976), operationalization of the construct is something of a guessing game and depends on the definition on which it is based. Some researchers propose specific sociolinguistic skills such as speech acts (Schmidt & Richards, 1980) and others have developed techniques for their elicitation (Grimm, 1975). Still others (Flavell et al., 1968; Krauss & Glucksberg, 1969; Legaretta, 1979; Politzer et al., 1981) have chosen referential communication tasks as indicators of communicative competence.

For the purposes of this investigation, the Canale (1981) framework of communicative competence guided inclusion of role-playing and story telling tasks as items in the CCT. These techniques have been used successfully to elicit natural-like communication from children (John et al., 1970; Legarreta, 1979; Overall, 1978; Bennett & Slaughter, 1981; Flavell et al., 1968; Politzer et al., 1981; Allen & Brown, 1976).

Role playing situations were created which would induce students to interact with real or hypothetical people. Story sequences were designed as stories of situations with which the students might identify. Successful interaction through both types of tasks is believed to indicate skill in the four areas (grammar, sociolinguistics, discourse and strategies) of communicative competence included in the Canale (1981) framework.

Role playing was performed using puppets (based on the technique of Overall, 1978). In the puppet interaction situations students were provided with brief descriptions of the setting (e.g., "This is Olivia; she is late for school; you take the part of Olivia and I will be the teacher"). Students were encouraged to interact freely with the tester. Although testers were provided with basic scripts (in order to insure some degree of consistency in administration) most of the interactions were free-form. The puppet interaction situations contained three home and three school items; domains (items) were alternated (see Appendix for descriptions of puppet interaction situations).

The story sequences were wordless stories of seven picture sequences each. Students were instructed to tell the story depicted by the pictures. Prior to the elicitation of the two sequences, students received shorter

(four pictures) trial items; the tester told the first story and the student told the next. Then students told the longer school and home stories. These were tape recorded. After each story telling, testers returned to five of the pictures and asked specific questions about the depicted dialogue in order to elicit particular speech acts (see Appendix for example of story sequence pictures).

Reliability and validity of the CCT were determined from both the field test (Sample #1) and study (Sample#3) data. This technical information is reported in Chapter Four.

Measures of Cognitive and Academic Functioning. Cartoon Conservation Scales (CCS). This test is a pictorial method for determining cognitive development which is unaffected by language background, SES or cultural experiences of the students (DeAvila & Pulos, 1979). It was developed to be group or individually administered using English, Spanish or any combination of the two languages. The test can be administered in any language.

> Because the CCS uses a pictorial format the approach has been particularly useful in assessing the intellectual development of children from language minority backgrounds. In this way, problems of direct translation have been avoided. (DeAvila, 1980b:1)

The CCS contains 8 subscales, each measuring a Piagetian concept. Each scale contains four items.

Cronbach's Alpha's range between .66 and .88 for third grade students (DeAvila, 1980a). Construct validity is reported from factor analytic studies in which distinct factors were found for each concept (Ulibarri, 1974; García et al., 1975). Criterion validity studies comparing concrete

Piagetian tasks with this pictorial mode report significant correlations for most tasks (DeAvila, 1980b). Furthermore, concurrent validity has been demonstrated by Ulibarri who found that the CCS and the Raven's Standard Progressive Matrices (A-E) (Raven, 1965) loaded on the same factor. Loadings on this factor by the CCS and Ravens were .71 for each test; his sample included 4th and 5th grade Hispanic students.

Comprehensive Test of Basic Skills (CTBS). The CTBS/S Level 1 English was administered in District #1 to students who were able to function in English. Other more limited English proficient students were given the CTBS Level 1 Spanish. District #4 gave only the English version.

The English test is an achievement battery; nine subtests were given (Table 5).

Table 5

CTBS/S Level 1

English

Subtest	# Items
Reading Vocabulary	40
Reading Comprehension	45
Spelling	44
Language Mechanics	20
Language Expression	30
Math Computation	48
Math Concepts	25
Math Applications	25
Reference Skills	20

Internal consistency is expressed in terms of Kuder-Richardson 20 correlations with the average coefficient being .92. Correlations with the Short Form Test of Academic Aptitude provided indications of concurrent validity; correlations with verbal IQ (as measured by the SFTAA) range from .37 to .78; non-language IQ correlations range from .43 to .67.

The Spanish test is also an achievement battery; only four of the subtests were administered. These parallel the subtests of the English form (Table 6).

Table 6

CTBS Level 1

Spanish

Subtest	# Items
Reading Vocabulary	40
Reading Comprehension	45
Math Computation	48
Math Concepts and Applications	50

Internal consistency is expressed in terms of Kuder-Richardson 20 correlations with most coefficients above .85. Correlations with the Spanish-English Reading and Vocabulary Screening (SERVS) instrument for third grade students average .62 (CTB/McGraw-Hill, 1979).

Metropolitan Achievement Tests (MAT). District #3 administered the Primary II, Form G, a battery of subtests containing the following:

Table 7

MAT Primary II, Form G

Subtest	# Items
Word Knowledge	40
Word Analysis	35
Reading	44
Spelling	30
Math Computation	33
Math Concepts	40
Math Problem Solving	35

The average split-half reliability coefficient for the subtests is .89. Validity data are reported as correlations with IQ tests for the levels above that of Primary II. The range of correlations for Elementary is .46 to .52; the more advanced levels report higher correlations (MAT, 1970).

Parent Questionnaire. All students in Sample #3 received copies of the Parent Questionnaire (written in Spanish). Ninety-six percent of the questionnaires were completed and returned. The researcher developed the ten-question form to ask parents the following information:

Place of birth of the student
Number of years student has lived in the U.S.
Number of siblings
Number of siblings living at home
Ages of siblings at home
Language(s) student uses with parents
Language(s) parent uses with student
Language(s) student uses with siblings
Mother's education
Father's education

Reliability and validity of the questionnaire items have not been determined. The information is susceptible to problems related to self-report data.

Data Collection Procedures

The investigation was conducted in four phases. Phase I included site preparation, gathering of teacher recommendations and parental permission, identification of Sample #1, preparation and acquisition of testing materials, tester training and pilot testing of the CCT. Phase II involved rating and scoring of Sample #1 data, selection of Sample #2 and testing of this sample on the BSM II. Phase III included selection of Sample #3, tester training and administration of a battery of tests to Sample #3. Phase IV concluded the study with a final debriefing session in which all previously selected samples were administered the CERAS Balance and students received their complimentary puppets. Reported below are only those procedures that involved data collection and scoring.

Student Identification. In order to identify potential students, meetings with all teachers were scheduled. Two group meetings were held. During each session a brief review of the study was provided; the research questions were not divulged. In this way it was hoped that teacher bias in sample selection or outcomes would be eliminated. Teachers received a short lecture on communicative competence. They were then asked to rate their Mexican-American non-monolingual students on a five-point Likert-type scale of communication skills, in English and Spanish. Only teachers who were bilingual could rate the Spanish proficiency of their students. Therefore, only 66% (N=154) of the possible sample was rated on communicative competence skills. Lists of all non-monolingual students were obtained from the 25 teachers. Once lists were drawn up, parent permission forms were sent home with students (N=386). Seventy-four percent of all students returned slips indicating parental permission to participate in the study.

Test Administrators. Fluent bilinguals, four Hispanics and one non-Hispanic raised in Latin America, were selected as testers. They were all female; two were graduate students in Bilingual Education at Stanford University; one was a bilingual preschool teacher from San José; one was a student at San Jose State University; one was a native bilingual with a master's degree in anthropology. All had previous living and working experience with Mexican-American children.

Testers were trained to administer the CCT and the BSM. Separate sessions were held to explain each test. Representatives from the Psychological Corporation trained testers on the BSM II (1980 edition). Three days later the researcher trained the same testers on the CCT. During these sessions testers practiced with materials, went through each tests' administration and received instructions on general testing procedures.

Field Testing. Twenty-one students from District #1 were individually administered the Communicative Competence Test and the BSM during four 30-minute sessions. Order of test language was counterbalanced; half of the sample received the BSM and CCT in English and then in Spanish and vice versa for the rest. The CCT was always administered first to establish rapport. Students were assigned randomly to condition. Within test item order for the CCT was constant; puppet interaction situations one through six were followed by the school and home story sequences. All testing was tape recorded for subsequent scoring and rating. Master tapes were made to ensure against missing data due to loss or damage of the original tapes.

Rater Training. Raters were chosen from teachers, aides and students in the Stanford and San José areas. Five fluent bilinguals (three

Hispanics and two non-Hispanics, two males and three females) were selected. Four sessions were held to train in rating and scoring procedures.

Rating sessions involved training on seven rating scales, discussions of the meaning of each scale level, listening to examplar tapes of the CCT (Session 1) and the BSM (Session 2) and rating according to the described levels of each scale or scoring according to the BSM II criteria.

The seven rating scales were based upon similar scales developed by Bartz (1974), Savignon (1972) the FSI (Clark, 1975) and Politzer et al. (1981). The scales included were:

Grammaticality	Effort to Communicate
Vocabulary	Amount of Relevant Information
Fluency	Global Speaking Proficiency
Comprehension	

The following is an example of the five levels taken from the Grammaticality Scale (see Appendix for complete listing of scales):

GRAMMATICALITY SCALE

Level 1. Student's use of grammatical structures is very limited; usually responds in single word utterances or with "Yes," "No," or "I don't know."

Level 2. Student attempts simple grammatical structures; these are characterized by many errors.

Level 3. Student attempts some complex grammatical structures; these are characterized by many errors; uses simple structures with some errors.

Level 4. Student's use of complex structures is better than level 3; some errors occur. Simple structures virtually errorless.

Level 5. Student uses complex grammatical structures with very few errors.

Training sessions continued until agreement on levels and their definitions were reached. Interrater reliabilities are reported in Chapter Four.

After training, raters were given six or seven of the field test tapes (Sample #1) and rating forms. They received instructions to listen to each tape no more than twice, to rate the Global Speaking Proficiency Scale, to put this rating scale out of sight and then to rate the other subscales.

Scorer Training. Scorer training was conducted after all ratings had been completed and returned; the same raters were trained to discrete point score BSM tapes. All scorers received copies of the BSM II Manual, Picture Booklets and practice Response Booklets. Procedures for scoring were explained; acceptable and non-acceptable answers were examined; rationales for scoring criteria were discussed; the central grammatical aspect of each answer was analyzed.

Following training, scorers were given the BSM response booklets which matched their tapes; booklets had not yet been scored by testers. Scorers scored the same tapes they had rated. Already familiar with the tapes, scorers would have little difficulty understanding them; scoring would, therefore, be less time consuming than if done by new scorers. Contamination of ratings with scorings was not likely; ratings had all been returned; moreover, discrete points, which are scored either 1 or 0, are little influenced by ratings. Interscorer reliabilities are reported in Chapter Four.

Samples #2 and #3. Sample #2 (N=185) was selected on the basis of dual language proficiency scores obtained from cumulative records in the schools. Students chosen were those who appeared to be either proficient

or limited in both English and Spanish. These students would be tested on the BSM II in order to locate those students who would currently be designated as LPB or PB.

Testers were to both test and score the BSM II. This procedure is recommended by the test developers. Therefore, testers were retrained on scoring procedures.

Tests were administered to Sample #2 in English and Spanish. Order of test language was counterbalanced with students randomly assigned to condition. Tests were scored immediately after administration. Due to transiency, 9 of the original 185 students could not be tested. Test responses were recorded on cassette tapes for future analysis.

English and Spanish total scores on the BSM II were crosstabulated in order to identify the 40 lowest (LPB) and 40 highest (PB) scoring students. Total scores rather than levels were used to select Sample #3. The levels did not discriminate the lowest scoring students from those in the middle.* Using Level 6 as the cutoff for proficiency, there were 108 students who scored as limited in both English and Spanish. Using Level 5 as the cutoff, there were only 9 students identified by the BSM II as LPB. Therefore, total scores on the BSM II were computed and plotted. On the BSM II total scores ranged from 0 to 21 (English) and from 1 to 21 (Spanish). Students scoring thirteen or below on both the English and Spanish tests were designated as LPB. Forty-three such students were found. Students scoring fourteen or above in English and fifteen or above in Spanish were designated as PB. Thirty-eight such students were found.

Sample #3 was then administered a battery of tests. The first was the CCS. This test can be group administered in Spanish, English or both

* See Sample #2 Crosstabulations in Appendix.

languages. In order to maintain cohesiveness during administration the students were divided into homogeneous proficiency groups within each school. If language heterogeneous students or only a few students were to be tested, testing was done individually.

After the CCS, students were individually administered the CCT and WNBD. Order of testing was as follows:

Session I:	CCT puppets	- Language A
	CCT stories	- Language A
Session II:	WNBD	- Language A
	WNBD	- Language B
Session III:	CCT puppets	- Language B
	CCT stories	- Language B

Within session order was maintained; language order was counterbalanced with students randomly assigned to condition. In a few cases, teachers did not want their students pulled out of class for three sessions. In these cases, two sessions, each starting with puppets (to establish rapport), were held. All tests were tape recorded. WNBD words were recorded and listed by testers on separate answer sheets.

Parent questionnaires were delivered to students along with pre-addressed, stamped envelopes for their return. Included were questions pertaining to length of residency, language use at home and parents' educational levels.

Three of the original five raters/scorers were retained for Sample #3. Tapes of the CCT for Sample #3 (N=81) were divided among the raters. Assignment to rater was done randomly within groups so that each rater had the same number of PB and LPB students. PB and LPB students were alternatively listed on the rating lists. Raters were instructed to go

down their lists, scoring each item (i.e., puppet interaction situation #1) for all students. After all ratings on one item were completed, the forms were put aside. Ratings of Item 2 were then made. Using this procedure, raters would be little influenced by a student's previous language sample.

Raters also scored a subsample of Sample #3 BSM tapes. The manual for the BSM II does not provide evidence of interscorer reliabilities. Because the BSM scores and CCT ratings were to be compared in the analyses, it was important to control for differences which might be due to scorer. In order to determine interscorer reliabilities among testers and raters, 40 tapes (twenty PB, twenty LPB) were randomly selected and distributed to raters. Moreover, the response booklets of these students were erased. Raters then scored student responses. Results of this analysis of interscorer reliability are reported in Chapter Four.

The CERAS Balance test was administered to Sample #3 in small groups within each school. Testers were trained to administer the test. Language order of testing was counterbalanced; students were assigned to condition based upon school site. Half of the students received the test first in English and then in Spanish and vice versa. Both tests were administered during a one hour session. This was done because of the previous time demands placed upon students; teachers were weary of pulling students out for any more sessions. After the CERAS Balance was administered, all students were thanked for participation in the project and given the puppet of their choice.

CHAPTER IV

RESULTS AND INTERPRETATION

Before answers to the four research questions were attempted, thorough examination and analysis of the technical properties of both the BSM II and CCT, the major linguistic and communicative tests used in this investigation, were performed. This chapter, therefore, includes discussion of these analyses as well as those addressing the research questions.

Bilingual Syntax Measure II: Reliability and Validity

Inter-scorer Reliability. In this research design, different language proficiency measures were scored by one of two sets of scorers-test administrators and outside scorers. In the case of the BSM, test administrators scored each test after it was given. Because results of various tests were to be compared in the analyses, it was important to determine whether any non-negligible effects arose due to scorer.

Interscorer reliabilities for the BSM II are not reported (Burt, Dulay, Hernández-Chávez & Taleporos, 1980). Therefore, a subsample of students (N=40) was randomly chosen from Sample #3 and their BSM tests were scored by both sets of scorers. Intraclass correlations (Table 8) indicate that for the purposes of this investigation, where comparisons of groups of students are to be made, interscorer reliabilities are acceptable.

Table 8

Intraclass Correlation Coefficients for BSM II

	R
BSM II TOTAL SCORE - Spanish	.81
BSM II LEVEL SCORE - Spanish	.79
BSM II TOTAL SCORE - English	.69
BSM II LEVEL SCORE - English	.65

N = 40

The relatively low reliabilities for the English level and total scores may have been the result of a restricted range of English proficiency in the Study sample. This is very likely to have been the case. Comparison of both BSM and CCT scores on the Field Test and Study samples (Tables 14-17) indicate that in the former, we find consistently higher interrater reliabilities for English on both tests. The Field Test sample was chosen from the extremes of the population (High in Spanish, Low in English and vice versa). The Study sample was comprised of a majority of limited English proficient students.

The lower English reliabilities suggest caution in using individual test results for diagnosis or placement, especially for students not yet proficient in English. Any interpretations made using either scores or levels should be made along with additional criteria. It is important to recognize that reliabilities of .65 and .69 indicate that there may be great fluctuation in a student's score if different testers were to administer the test.

Order and Tester Effects. Table 9 displays the results of analyses of variance to determine effects of test order (English - Spanish vs.

Spanish - English) and tester.

Table 9

Analyses of Variance Showing BSM Tester and Order Effects (Sample #2)

Source	N	Mean	SD	df	Sum of Squares	F
BSM English Order	174	13.70	3.47	1	17.0043	1.42
BSM English Tester				4	343.5028	7.15***
BSM Spanish Order	166	12.05	4.60	1	79.6697	3.73
BSM Spanish Tester				4	641.7849	7.51***

***p < .001

As indicated, the effect of order was not statistically significant. This is probably due to the counterbalanced design. Tester, however, does show main effects in both English and Spanish. Such strong effects (significant at the .001 level) may have been due to the confound of tester with school/examinee. Testers, for the most part, did not rotate among the 18 schools. Therefore, the tester effect is confounded with systematic differences in the ability of students at different schools. The strength of the effect, although confounded with schools, indicates that additional caution in use of the BSM is warranted. The test may be sensitive to differences in administrators.

Internal Consistency. Split-half reliabilities were calculated for the BSM II, English and Spanish. Each test was divided in half, with every other item assigned to one or the other half. Pearson correlations between halves were calculated. The coefficients were then corrected

with the Spearman-Brown Prophecy Formula in order to compensate for the reliability estimated from a test only one-half the length of the original form. These relabilities are reported in Table 10.

Table 10

Split Half Reliabilities of BSM II (Sample #2)

Test	Pearson's R	Corrected R
BSM II Spanish	.79	.88
BSM II English	.62	.77

N = 174

The BSM II Technical Manual (Burt, Dulay et al., 1980) reports slightly larger internal consistency reliability coefficients. Alpha coefficients for English totals (N = 500) and Spanish totals (N = 401) are reported as .90 and .82 respectively. These differences may be due to small sample size and restricted range of English and Spanish abilities in this third grade sample.

BSM Levels vs. BSM Total Scores. Level designations for the BSM II range from 1 to 6 with 6 as the highest level. These levels have been assigned on the basis of criteria established during the test's original field testing. The criteria are based upon an ordering-theoretic method which groups linguistic phenomena into levels of implicational relationships. There are "no strict empirical guidelines" (Burt, Dulay et al., 1980, p. 21) for determination of item orders or cutoff levels. The test's authors, therefore, imposed their own guidelines to designate which items confirmed their theoretical ordering and where cutoffs were to be drawn.

Having determined which items are theoretically more difficult (acquired later) than others, items were assigned to levels. The first (easier) portion of the hierarchy was used to determine if a student was in the Level 3/4 group. The latter portion of the hierarchy was used to determine if a student was in the Level 5/6 group. Level 1 is assigned if the student answers fewer than three questions out of the first four in any language; Level 2 is assigned if the student answers fewer than thirteen questions in the appropriate language. The difference between Levels 3 and 4 or between Levels 5 and 6 is determined by the number of correct items within the level groups. The differences in these levels are not determined by answering correctly more difficult items within each group.

Thus, the procedures for determining language proficiency designations result in level designations not monotonically related to the total number correct. The BSM II (like the BSM I) can assign a level of Proficiency to a student who has a lower total score than a student assigned to the Limited levels. This characteristic calls into question the validity of the test's levels to indicate proficiency. Would a simple total score be a more valid indicator of language proficiency?

In order to evaluate the validity of these levels, the levels and total scores were correlated (1) with other tests of language proficiency, (2) with achievement, and (3) with each other. Results are listed in Tables 11 and 12. The correlation of BSM English Total score with every other variable exceeded that of BSM English Level with the same variable.

Table 11

Intercorrelations of English BSM II With Measures of Language Proficiency and Achievement (Sample #3)

Test	BSM E Level	BSM E Total Score
CKRAS English (N = 77)	.28 *	.39 ***
CCT English (N = 75)	.05	.22
WNBD English (N = 77)	.07	.31 **
CTBS English Reading Total (N = 30)	.22	.35
CTBS English Language Total (N = 28)	.24	.33
CTBS English Math Total (N = 28)	.34	.44 *
CTBS English Total Battery (N=28)	.29	.40 *
MAT Total Reading (N = 33)	.17	.23
MAT Total Math (N = 34)	.14	.46 **

* $p < .05$
** $p < .01$
*** $p < .001$

Similar analysis was performed for the Spanish levels and Total Scores. Because only 9 students had achievement in Spanish data, correlations are reported with other measures of oral/aural language proficiency, only.

Table 12

Intercorrelations of Spanish BSM II
With Measures of Language Proficiency (Sample #3)

Test	BSM S Level	BSM S Total Score
CERAS Spanish (N = 77)	.19	.32 **
CCT Spanish (N = 74)	.35 **	.43 ***
WNBD Spanish (N = 75)	.44 ***	.55 ***

** $p < .01$
*** $p < .001$

Table 13

Correlations Between BSM Total Scores
and Levels (Sample #3)

Language	Correlation Coefficient
English	.80 ***
Spanish	.87 ***

*** $p < .001$
N = 77

These correlations indicate to what extent the BSM II Levels and Total Scores are measuring something in common with other measures, presumably, language proficiency. The consistently lower correlations of the Levels with all criterion variables as well as the correlations between BSM Levels and Scores suggest that the Levels are less reliable than the Total Scores. The squares of the English and Spanish correlations in Table 13 are .64 and .76 respectively. These coefficients

indicate that the BSM Levels could explain only 64% and 76% of the variance in total scores. A certain amount of precision has been lost in collapsing the Total Scores to a six point scale and, in addition, the method of determining level designations has led to diminished reliability. Using the Levels, alone, one loses information about the language proficiency of students.

The lack of exact correspondence between Levels and Total Scores as well as the low correlations between Levels and other language and achievement measures may be due to the way in which the levels are determined. Not only are there no strict empirical guidelines for determination of the "true" hierarchy, but the designation of Level 3 vs. Level 4 or Level 5 vs. Level 6 is not determined by passing progressively more difficult items. Rather, these Level designations are determined by the number of items correctly answered within each grouping. That is, levels are determined by the number of correct answers within each group rather than Levels 4 and 6 indicating passage of more difficult items.

Additionally, the responses elicited by the BSM and scored as correct or incorrect vary from a single response to three possible correct responses per item. These elicited responses may be one of three distinctly different grammatical structures, each of which may be theoretically more or less difficult than the construction for which the item was ordered. Actually, 27% of the English items and 24% of the Spanish items can be scored correctly with two or three possible answers.

The validity of the BSM II Levels has been questioned due to results of above mentioned analyses. It is recommended that Total Scores be used rather than Levels for determining placement of students from non-English language proficiency backgrounds.

Communicative Competence Test:
Reliability and Validity

Inter-rater Reliabilities. The primary purpose of field testing the CCT was to estimate reliability of the instrument. Inter-rater reliabilities for total ratings were approximated using intraclass correlations based on the assumption that scorers were drawn at random from the population of potential scorers. Table 14 displays these reliabilities.

Table 14

Sample #1 Intraclass Correlations of CCT
Field Test
(Five Raters)

	r
CCT Total Ratings - Spanish	.92
CCT Total Ratings - English	.34

In order to increase the interrater reliability of the English total ratings on the CCT, one of the raters was dismissed. Reanalysis of variance components and recomputations of intraclass correlations reveals that the reliability estimate was improved considerably for English total ratings.

Table 15

Sample #1 Intraclass Correlations of CCT
Field Test
(Four Raters)

	r
CCT Total Ratings - Spanish	.99
CCT Total Ratings - English	.67

For Sample #3, each of the six puppet interaction situations and two wordless story sequences was rated on each of the seven rating subscales. This procedure differed from rating of the field test sample (Sample #1); for the field test sample there was one set of ratings for the entire test.

Because new rating procedures were used, new interrater reliabilities were determined. Three raters each rated eighteen tapes in both English and Spanish. Then, generalizability coefficients (Cronbach Alphas) were computed. Table 16 lists the results of this computation.

Table 16

Generalizability Coefficients for CCT

Sample #3 (N = 77)

English	Spanish	Generalizability
.88	.98	average of the same 3 raters for all subjects
.71	.95	one rating per subject, same rater for all subjects
.39	.80	one rating per subject, by rater randomly chosen from infinite population
.46*	.84*	one of a universe of 3 raters rating each subject

*Appropriate for this study.

The noticeable differences between English and Spanish coefficients may have been attributable to lack of sufficient rater training, rater differences, student differences or many other plausible factors. If training were the important factor, Spanish coefficients should be equally affected. Training was, after all, provided for both language forms during the same sessions. Careful examination of the coefficients

indicates that one rater was a bit more conservative than the other two; however, this does not account for most of the difference between Spanish and English coefficients. The conclusion, which is supported by the data (see Table 17) is that the range of student's language proficiencies in English is more restricted than in Spanish. Through examination of score distributions on CCT, BSM and CERAS, we can see that the English language proficiencies of these students fall within a narrower band than for Spanish.

Table 17

Score Distributions for BSM, CERAS, CCT and WNBD

Language	Test	Mean	SD	Min. Score	Max.Score
English	BSM - Total Score (N = 174)	13.7	3.47	3	21
Spanish	BSM - Total Score (N = 166)	12.1	5.01	0	21
English	CERAS - Total Score (N = 190)	17.4	2.6	5	20
Spanish	CERAS - Total Score (N = 190)	17.3	2.8	4	20
English	CCT - Total Score (N = 78)	236.3	28.81	149	279
Spanish	CCT - Total Score (N = 78)	241.1	41.82	40	280
English	WNBD - Total Count (N = 80)	64.0	27.27	28	142
Spanish	WNBD - Total Count (N = 78)	57.2	24.82	6	113

With the exception of word naming, the language proficiency measures show that English scores vary less than Spanish scores. Intraclass correlations from the Sample #1 provide further support for this explanation. Intraclass correlations for 4 raters' ratings of the CCT (Table 15) was .67. The higher coefficient for the Sample #1 than for Sample #3 may be due to the wide range (very high or very low) of English and Spanish language proficiencies among that Sample (see Sample Selection).

In sum, similar conclusions and recommendations for individual and group comparisons using the BSM and CCT are warranted. For use in this study and for making group comparisons, all reliability coefficients are judged as acceptable. If individual diagnosis or placement decisions are to be made using the rating system as it exists, three caveats should be applied:

Caveat #1: try to have more than one person rate the same student or have one trained rater rate all children;

Caveat #2: do not try to distinguish among students very similar in language ability;

Caveat #3: use other indicators in conjunction with the CCT or BSM.

The population in question is learning English as a second language; it would make sense, therefore, that English language proficiency scores are less varied, more restricted, than those of Spanish. The reliability of the instrument, if we can generalize from Spanish ratings, would improve considerably if a normal distribution of students had been used.

Order and Tester Effects. There were no main effects of either test order or tester for the CCT. Table 18 indicates the results of analysis of variance to determine order and tester effects.

Table 18

Analyses of Variance Showing CCT Tester and Order Effects (Sample #3)

Source	N	Mean	SD	df	Sum of Squares	F
CCT English Order	78	236.32	28.81	1	335.19	.40
CCT English Tester				4	5715.34	1.72
CCT Spanish Order	78	241.05	41.80	1	2261.92	1.29
CCT Spanish Tester				4	6604.05	.95

The lack of order effect indicates that the English and Spanish testings were independent of one another. Prior testing in the other language did not affect average performance. Lack of tester effect indicates that, although testers were nested in schools, systematic differences among testers were not important. This differs from the BSM; that test is apparently more sensitive to language proficiency or other examinee/tester variables.

Internal Consistency. Internal consistency of the CCT is provided through split-half reliability coefficients and inter-subscale correlations coefficients (Tables 19 - 23).

Table 19

Split-Half Reliabilities of CCT (Sample #3)

Test	Pearson's R	Corrected R
CCT Spanish	.82	.90
CCT English	.85	.92

N = 78

Split-half reliability coefficients were determined through Pearson Product Moment Correlations with the Spearman-Brown Prophecy correction.

Inter-subscale reliability coefficients are provided in terms of Pearson Product Moment Correlations.

Table 20

CCT Subscale Correlation Coefficients

Sample #1 - English

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Global Speaking Proficiency Proficiency	1.00						
(2) Grammaticality	.91	1.00					
(3) Vocabulary	.88	.87	1.00				
(4) Fluency	.91	.86	.94	1.00			
(5) Comprehension	.79	.71	.75	.86	1.00		
(6) Amount of Relevant Communications	.78	.66	.78	.86	.86	1.00	
(7) Effort to Communicate	.54	.46	.59	.66	.81	.84	1.00

N=21

Table 21

CCT Subscale Correlation Coefficients

Sample #3 - English

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Global Speaking Proficiency	1.00						
(2) Grammaticality	.94	1.00					
(3) Vocabulary	.96	.91	1.00				
(4) Fluency	.94	.85	.94	1.00			
(5) Comprehension	.73	.71	.70	.75	1.00		
(6) Amount of Relevant Communications	.66	.54	.66	.70	.74	1.00	
(7) Effort to Communicate	.84	.76	.80	.82	.73	.76	1.00

N=77

Table 22

CCT Subscale Correlation Coefficients

Sample #1 - Spanish

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Global Speaking Proficiency	1.00						
(2) Grammaticality	.77	1.00					
(3) Vocabulary	.78	.82	1.00				
(4) Fluency	.83	.87	.85	1.00			
(5) Comprehension	.43	.53	.55	.49	1.00		
(6) Amount of Relevant Communication	.79	.77	.83	.91	.65	1.00	
(7) Effort to Communicate	.88	.79	.79	.86	.85	.86	1.00

N=21

Table 23

CCT Subscales Correlation Coefficients

Sample #3 - Spanish

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Global Speaking Proficiency	1.00						
(2) Grammaticality	.95	1.00					
(3) Vocabulary	.97	.97	1.00				
(4) Fluency	.94	.94	.96	1.00			
(5) Comprehension	.73	.74	.76	.80	1.00		
(6) Amount of Relevant Communication	.75	.75	.76	.73	.79	1.00	
(7) Effort to Communicate	.68	.69	.71	.67	.73	.70	1.00

N=77

The internal consistency reliability of the CCT is stable and highly uniform for Global Speaking Proficiency, Grammaticality, Vocabulary and Fluency subscales but appears to fluctuate on Comprehension, Amount of Relevant Communication, and Effort to Communicate. Inter-subscale reliabilities are acceptable for comparisons of groups, as is the case with this investigation. Use for individual diagnosis or placement could be made with substantial accuracy if the subscales of Comprehension, Amount of Relevant Communication and Effort to Communicate were not used. It appears as though these three subscales may be more difficult to rate because non-verbal, paralinguistic cues are not available to raters listening to tape recordings. If these scales were used after face-to-face interaction with students, perhaps higher intercorrelations with other subscales would be obtained.

The reliabilities for both the BSM II and the CCT have been determined and accepted as sufficient for use in this investigation. These two tests are the major instruments used in the analyses. Attention is now turned to answering the four research questions.

Question #1 Among students classified as "limited proficient bilingual," is the total repertoire of grammatical skills distributed across languages? If so, is this distribution complementary?

The rationale for the complementary distribution of language skills is based on Ferguson's (1964) notion of diglossia. In some bilingual communities, languages are separated for functionally differentiated purposes. According to Fishman (1971), no society needs two languages for one and the same set of functions. It may be the case that the students

in San José live in a diglossic environment in which they separate the use of English and Spanish for specific purposes; if so, they may also distribute the grammatical skills functionally. If the students are tested in grammar in either language the results might indicate limited language; if tested in both, perhaps the distribution of grammatical proficiencies would indicate proficient status.

The hypothesis underlying the notion of complementary distribution is the following: Some students acquire skill X in English, some acquire it in Spanish; once they demonstrate the skill in one language, they are less likely to demonstrate it in the other language.

In order to test this hypothesis, English and Spanish semantically equivalent responses to items on the BSM and CERAS Balance were chosen and scores crosstabulated. Only seven of the BSM items could be reliably categorized as semantically equivalent because of item design and scoring procedures. Although similar questions and pictures are used in both English and Spanish forms, the responses elicited vary across language forms and within each item. That is, Question #3 in English is not semantically equivalent to Question #3 in Spanish; it is not necessarily equivalent to any other item on the Spanish form. Moreover, where semantically equivalent responses could be scored, other responses could also have been scored as correct. For approximately one-fourth of the test, a correct score does not necessarily indicate that the response is the semantically equivalent one. Therefore, only 7 items from both English and Spanish forms could be chosen as equivalent; for these items, there is only one possible correct response. The CERAS Balance does not suffer from these problems. The Spanish form is a translation of the English. Therefore, there is a direct semantic equivalence across all CERAS items.

The sample used in analyses was composed of those students who scored as limited in both English and Spanish on the BSM II (levels 1-5). Approximately 50 students met these criteria.

Crosstabulations of correct and incorrect responses for all item pairs were performed and Phi Coefficients obtained. Given the underlying hypothesis, a negative coefficient would lend support for this hypothesis. As can be seen in Table 24, no significant negative Phi Coefficients were obtained.

Table 24

Summary of Distributions of Semantically Equivalent Syntactic Items (N=50)

SKILL	BSM/CREAS English Item #	BSM/CREAS Spanish Item #	% demonstrating skill in ENGLISH ONLY	% demonstrating skill in SPANISH ONLY	% demonstrating skill in BOTH	% demonstrating skill in NEITHER	Phi Coefficient	Probability
BSM IX								
conditional	11	18	25.5	19.6	11.8	43.1	.003	.9305
conditional	11	9	25.5	25.5	10	39	-.116	.4048
word order in questions	15	5	13.4	17.3	40.4	28.9	.335	.0158*
word order in questions	15	11	18	6	38	38	.542	.0001*
word order in questions	15	20	19.2	13.5	36.5	30.8	.348	.0130*
conditional sum + participle	17	4	2	21.6	3.9	72.5	.236	.0916
irregular past	20	3	28	16	24	32	.131	.3592
CREAS Salomon								
singular to plural conversion	2	2	9.6	5.8	84.6	0	-.081	.5606
singular to plural conversion	3	3	13.4	1.9	82.7	0	-.080	.6468
plural to singular conversion	5	5	7.7	3.85	84.6	3.85	.348	.0122*
plural to singular conversion	6	6	11.5	13.5	63.4	9.6	.273	.0473*
present to past conversion	8	8	13.5	23.1	59.6	3.8	-.044	.7264
present to past conversion	9	9	5.8	32.7	61.5	0	-.172	.2137
past affirmative to past negative	11	11	25	3.8	6.92	1.9	.036	.7963
past affirmative to past negative	12	12	5.8	5.8	86.5	1.9	.188	.1763
location	14	14	11.5	5.8	78.9	3.8	.223	.1086
location	15	15	13.5	0	84.6	1.9	.328	.0179*
question: indirect to direct	17	17	7.7	1.9	90.4	0	-.040	.7707
question: indirect to direct	18	18	11.5	5.8	82.7	0	-.089	.5193
command: indirect to direct	20	20	3.8	5.8	90.4	0	-.049	.7212
command: indirect to direct	21	21	5.8	7.7	86.5	-	-.071	.6063
question: direct to indirect	23	23	3.85	3.85	90.4	1.9	.293	.0349*
question: direct to indirect	24	24	3.85	13.5	78.8	3.85	.249	.0720
command: direct to indirect	26	26	38	15.4	80.8	0	-.085	.5386
command: direct to indirect	27	27	17.3	5.8	71.2	5.8	.231	.0961
comparatives	29	29	13.5	19.2	67.3	0	-.192	.1632
comparatives	30	30	19.2	19.2	44.2	17.3	.171	.2185

* p <.05 level

There were, however, significant positive Phi Coefficients in 7 of the 27 item pairs. These results indicate that the hypothesis of distribution of grammar across languages is clearly not supported. They also indicate support for previous research on bilingual language acquisition by Kessler (1971). In her work with Italian/English first and second grade bilinguals she found that structures shared by the two languages were acquired "in the same sequential pattern and at approximately the same rate." The theoretical support for this simultaneous acquisition of grammatical structures comes from transformational grammar. It is assumed that these semantically equivalent structures "share elements of processes in their deep or underlying structure." Results of this analysis indicate support for this language universal hypothesis.

The notion of complementary distribution of language skills may make more sense in terms of vocabulary which is functionally differentiated by domains. Indeed, most of the research on diglossia uses vocabulary measures as indicators. Therefore, Question #2 asks if vocabulary and other communication skills (which can be differentiated by domain) are distributed, within this population, across English and Spanish.

Question #2 Among students classified as "limited proficient bilingual," is the total repertoire of vocabulary and communication skills distributed across languages as a function of domain?

In order to answer this research question, two-way analyses of variance with repeated measures were performed on the English and Spanish home and school sums for each of WNBD, CCT combined scales and separate CCT subscales. In these analyses, a main effect of domain or a main

effect of language would indicate systematic differences in proficiency, or possibly non-equivalence of test forms. An interaction of language by domain would support the hypothesis of different language skill distributions across domains.

Results of these analyses indicate that (1) students are better in Spanish than English, and (2) with only one exception, there is no significant main effect of domain. Table 26 provides a summary of the interaction effects from the ANOVAs, which test the hypothesis underlying Question #2. Table 25 provides a summary of main effects due to language or domain.

Table 25

Repeated Measures Analysis of Variance - Main Effects

Sample #3 (1,76 df) N = 78

Test	Language	Domain	Mean	SD	F-Domain Effect	F-Language Effect
WRAD	English	Home	35.17	15.19	66.38 ***	9.61 **
	English	School	29.68	13.76		
	Spanish	Home	32.27	14.00		
	Spanish	School	25.81	12.05		
CCT Combined scales	English	Home	118.97	13.63	2.0	5.35 *
	English	School	118.10	16.44		
	Spanish	Home	122.44	16.22		
	Spanish	School	123.93	15.65		
CCT Vocabulary	English	Home	16.47	2.73	.62	6.25 *
	English	School	16.34	3.08		
	Spanish	Home	17.46	3.29		
	Spanish	School	17.58	3.19		
CCT Grammaticality	English	Home	16.96	2.20	.36	6.40 *
	English	School	16.71	3.62		
	Spanish	Home	17.80	3.07		
	Spanish	School	17.91	3.15		
CCT Comprehension	English	Home	18.86	1.38	.53	1.69
	English	School	18.80	1.95		
	Spanish	Home	19.00	1.88		
	Spanish	School	19.22	1.50		
CCT Fluency	English	Home	17.11	2.15	1.16	8.81 **
	English	School	16.96	2.34		
	Spanish	Home	18.13	2.87		
	Spanish	School	18.03	3.02		
CCT Amount	English	Home	16.79	2.45	2.69	.17
	English	School	16.55	2.60		
	Spanish	Home	16.22	2.70		
	Spanish	School	16.87	2.84		
CCT Effot	English	Home	16.12	2.03	2.31	.14
	English	School	16.12	2.30		
	Spanish	Home	16.03	2.47		
	Spanish	School	16.38	2.48		
CCT Global speaking proficiency	English	Home	16.41	2.50	.53	3.76
	English	School	16.29	2.78		
	Spanish	Home	17.16	3.29		
	Spanish	School	17.09	3.51		

* $p < .05$ ** $p < .01$ *** $p < .001$

Table 26

Repeated Measures Analyses of Variance - Interaction Effects
Language and Domain Variables (Sample #3)

Measure	Source	df	Sum of Square	Mean Squares	F
WNBD	language X domain	1	18.26	18.26	.89
	error	76	1550.99	20.41	
CCT combined subscales	language X domain	1	102.50	102.50	2.32
	error	76	3186.75	44.26	
CCT vocabulary	language X domain	1	.05	.05	.05
	error	76	81.45	1.09	
CCT grammaticality	language X domain	1	2.40	2.40	1.83
	error	76	98.35	1.31	
CCT comprehension	language X domain	1	1.45	1.45	1.17
	error	76	93.30	1.24	
CCT fluency	language X domain	1	.03	.03	.02
	error	76	115.22	1.54	
CCT amount	language X domain	1	14.77	14.77	8.24 **
	error	76	134.48	1.79	
CCT effort	language X domain	1	2.40	2.40	2.40
	error	76	74.85	1.00	
CCT global speaking	language X domain	1	.05	.05	.04
	error	76	91.45	1.22	

**p < .01

N = 78

There is little support for the hypothesis that students in this population may have been raised in diglossic environments, using English for one set of purposes (i.e., in school) and Spanish for another (i.e., in the home). This result supports the work of Beltramo (1972) who concluded from his research in the same area, that the bilingualism of Mexican-Americans in San Jose is not generally characterized as diglossic. Only for the WNBD tasks was domain a significant main effect as indicated in Table 25. Students appear to know more home than school vocabulary.

There is considerable evidence that students were more proficient in Spanish than in English. Except for WNBD, in all other cases where significant main effects of language were found, the stronger language was Spanish. The result contradicts the findings of Dulay and Burt (1980) who report that most of the limited proficienct bilinguals in San Jose were limited-English dominant.

Neither domain nor language effects were found for Comprehension, Amount of Relevant Information or Effort to Communicate. As indicated by the inter-subscale correlations (Tables 20-23), these three scales may be more difficult to rate, less reliable and, hence, less susceptible to main effects.

Domain does not interact with language except in the Amount of Relevant Information (Table 26). This result may be due to the unreliability of this scale or to differences among students. Examinations of means and standard deviations for this scale indicate that scores are very close. Therefore, the significant interaction effect is probably not attributable to student differences (Table 25).

The results of WNBD analyses are quite different from results of the CCT. For WNBD, students appear to be better in English than in Spanish

and better in the home than the school domain. Explanations for these contradictory findings may be that vocabulary naming is easier in English and that the home picture is easier than the school picture. Perhaps third grade students are used to word naming from their ESL classes. Moreover, the two pictures used in the word naming tasks are quite different; the home is divided into four neat, uncluttered rooms; the school is one room of chaos! (See Appendix for pictures.)

Question #3 Among students classified as "limited proficient bilingual," is there a difference in assigned bilingual proficiencies when different types of proficiency tests are used to assess verbal behavior?

In order to assign levels of monolingual/bilingual proficiency, cut-off points must be determined which differentiate, at least minimally, between those who are proficient in the language(s) and those who are not. These levels have been determined for the BSM by the test developers. The other measures—CCT, CERAS Balance and WNBD—do not have level designations. Therefore, to answer this research question, procedures for producing level designations (limited vs. proficient) on all language proficiency tests were designed. Two procedures were used. The first is that of Equivalent Percentile Cut-offs (Thorndike, 1971; Ulibarri et al., 1981) based upon percent of the sample who scored as Proficient, in English and Spanish, on the BSM II. The second procedure, Grade-level Regression Cut-offs, should have more predictive validity; levels were determined on the basis of each test's regression on achievement (CTBS English, Total Battery). Proficiency was the score which predicted grade-level achievement; limited were scores below this cut-off.

Analysis of Concurrence (crosstabulations) between level designations for tests were performed; crosstabulations between proficient and limited language proficiency designations were performed using both the Equivalent Percentile and Grade-level Regression cut-offs.

Three other procedures for indicating agreement/disagreement among language proficiency measures were undertaken. Correlations among all language proficiency tests were performed and resultant coefficients examined. Multiple regressions of achievement on language proficiency and Factor Analyses were undertaken.

Equivalent Percentile Cut-off Procedure. The Equivalent Percentile Cut-off Procedure provides a means to equalize tests which do not have cut-offs for Limited and Proficient levels to match the cut-offs on a test with these levels. In this research, only the BSM II has level designations. Although the validity of these levels has been seriously questioned (Ulibarri, Spencer & Rivas, 1981; Gillmore &Dickerson, 1979; Cervantes & Nakano, 1979) they are, in fact, used by school districts and language researchers to measure and diagnose language proficiency and to prescribe language programs for non-English language background students. These levels are accepted by state and federal education agencies for use in identifying such students who have limited language proficiencies in English and Spanish. Descriptions of the BSM II Levels are listed in the Appendix.

In these descriptions, Level 5 is labeled Proficient I; however, the cut-off for proficiency can vary. Either Level 5 or Level 6 may be chosen as the cut-off for proficiency. Level 5 can mean either limited or proficient in language skills. California's State Department of

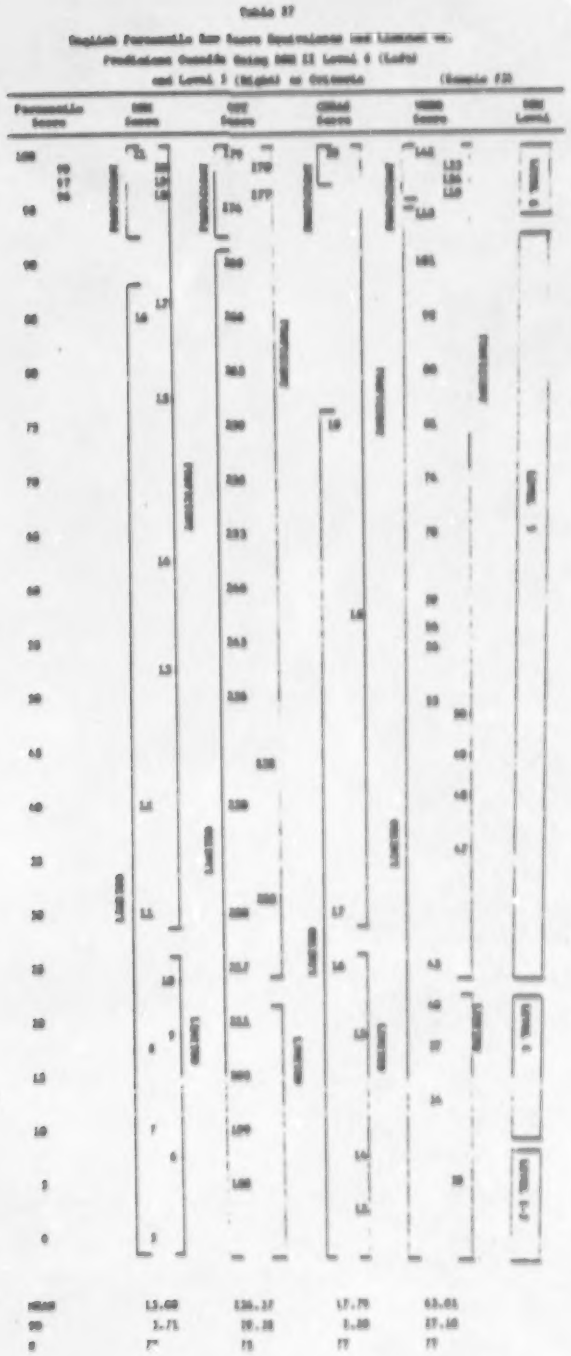
Education Office of Bilingual-Bicultural Education designates Level 6 as the cut-off for proficiency. Moreover, the authors of the test state that "English-speaking majority students usually score in the top level of these tests; thus, students scoring in the top level are considered proficient while those who score in the lower levels are usually considered limited to varying degrees" (Dulay & Burt, 1980).

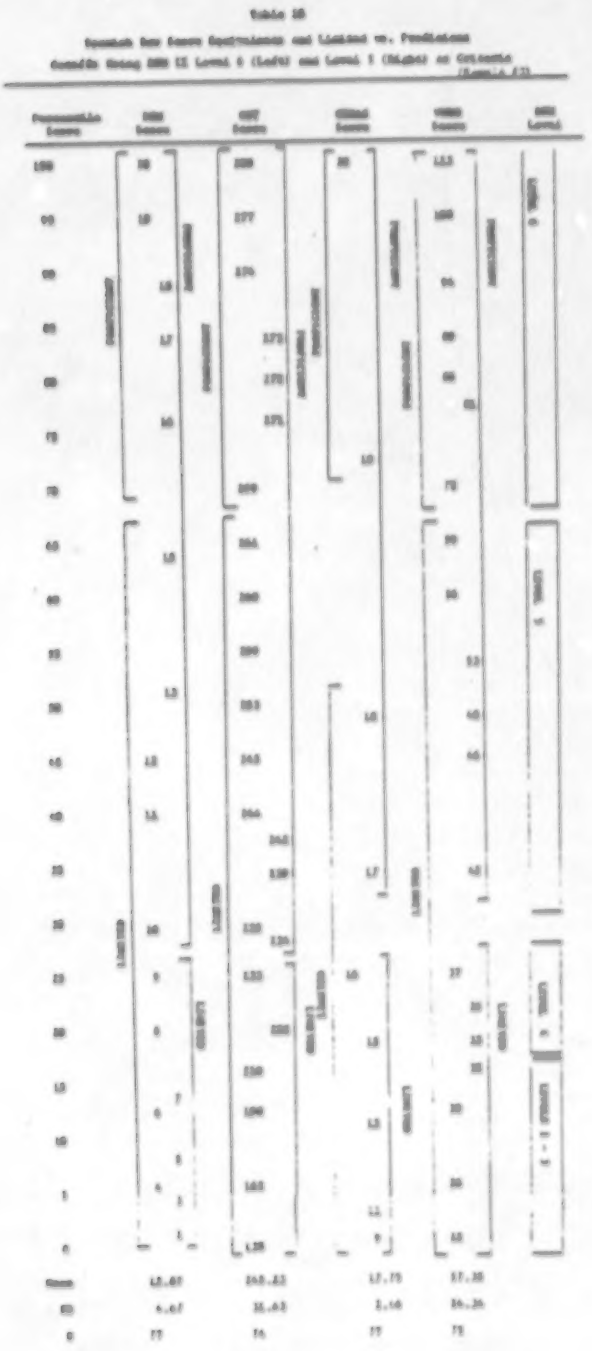
However, the authors also provide justification for the use of Level 5 as proficient (Burt, Dulay, Hernández-Chávez & Taleporos, 1980):

> While most native-English-speaking students score at Level 6, 44.4% of them were assigned a Level 5 score. It has been the experience of the BSM II authors that the English proficiency of some native-English-speaking students in some school districts is below that of other districts. Perhaps educational placement in these districts should be made with reference to the ability level of the local native-English-speaking group; special education programs may be considered to be appropriate only for students who score below their district norms. (p. 22)

Because two different criteria can be applied for assignment of Proficient vs. Limited Level designations using the BSM II, analyses using the Equivalent Percentile Cutoff procedure were performed with both Level 6 and Level 5 as the cutoffs for proficiency.

Using Level 6 as Proficient and Levels 1-5 as Limited, percentages of students who scored Level 6 in English and Spanish were used. Five percent (N=4) of the students in Sample #3 scored Level 6 in English; 32.5% (N=26) scored Level 6 in Spanish. Scores on the CCT, CERAS and WNBD were equated with the BSM Levels at the points on each test above which the top 5% (in English) and 32.5% (in Spanish) of the students fell. As can be seen in Tables 27-28, the cutoffs were made as close as possible to the BSM cutoffs.





Having established these levels on all language tests, the first analyses to answer Question #3 were performed. Crosstabulations of language proficiency levels on each test were made, crossing pairs of tests. Crosstabulations were made within each language. The results of these analyses reflect the percent of agreement, disagreement, and degree of association between each pair of tests (Chi square).

Table 29

Equivalent Percentile Procedure
Crosstabulations of Limited/Proficient
Level Designations - Proficient = Level 6 (Sample #3)

Tests	Percent Agreement	Percent Disagreement	Chi Square
English			
BSM with CCT	90	10	.222
BSM with CERAS	76	24	1.60
BSM with WNBD	93	7	3.546
CCT with CERAS	76	24	1.60
CCT with WNBD	90	10	.222
CERAS with WNBD	76	24	1.60
Spanish			
BSM with CCT	70	30	7.337 **
BSM with CERAS	53	47	.228
BSM with WNBD	71	29	.154
CCT with CERAS	53	47	.238
CCT with WNBD	71	29	8.381 **
CERAS with WNBD	59	41	2.851

N = 75

** $p < .01$

Lack of significant Chi Squares indicates that for the English measures based upon Equivalent Percentile cutoffs anchored on the BSM II levels, with proficient as Level 6, there is little association between measures that could not be accounted for by chance. To look only at the percent of agreement would be misleading; high percent of agreement is expected because 95% of the students are in one of the categories (limited). Thus, the language measures concurred on levels of students no better than, and often worse than, chance. Moreover, the range of disagreement indicates that even when equivalent percentile cutoffs are used, between 7% and 47% of the students are inconsistently classified.

The two significant Chi Squares among Spanish measures indicate that although approximately 30 percent of the students were classified differently, there is significant association between CCT Spanish and BSM Spanish and between CCT Spanish and WNBD Spanish. The lack of significant association between BSM Spanish and WNBD Spanish indicates that the CCT may be the common measures. That is, communicative competence appears to share some things in common with both grammar and vocabulary.

In order to investigate the agreement among instruments on students' bilingual classifications, comparisons of the joint classification of students' English and Spanish proficiencies on each instrument were performed. The joint classifications of interest were limited-proficient bilingual (LPB) and non-limited proficient bilingual (non-LPB).

Relative language proficiency (RLP) or bilingual classification, is the degree to which proficiency is exhibited in both L1 and L2. Table 30 displays RLP possibilities. For the purposes of this investigation (in order to have sufficient N per cell) the four RLP categories were collapsed

into two: LPB (limited proficient bilingual) and non-LPB. If proficiency was attained in one or both languages (according to the Equivalent Percentile Procedure) a student was classified as non-LPB.

Table 30

Bilingual Classifications (RLP)

L1 Proficiency	L2 Proficiency	RLP Categories	Bilingual Classifications
limited	limited	limited proficient bilingual (LPB)	LPB
limited	proficient	L2 dominant	non-LPB
proficient	limited	L1 dominant	non-LPB
proficient	proficient	proficient bilingual (PB)	non-LPB

Crosstabulations of bilingual classifications for each pair of measures were performed. Results indicate that, although the Equivalent Percentile Procedure (using Level 6 as proficient) provides some association between measures, the percent of students consistently classified as LPB or non-LPB is, on the average, only 61.5 (see Table 31).

Table 31

Results of Crosstabulations Between Bilingual Classifications

Proficient=Level 6 (Sample #3)

Tests	Percent Agreement	Percent Disagreement	Chi Square
BSM with CCT	66	34	4.619 *
BSM with CERAS	52	48	.812
BSM with WNBD	71	29	9.990 **
CCT with CERAS	54	46	1.641
CCT with WNBD	70	30	7.288 **
CERAS with WNBD	56	44	3.129

N = 75
*p < .05
**p < .01

Similar analyses were performed using Level 5 as the cutoff for proficiency. Results of crosstabulations between tests based upon the Equivalent Percentile Cutoff procedure are listed in Tables 32 and 33.

Table 32

Equivalent Percentile Procedure
Crosstabulations of Limited/Proficient
Level Designations Proficient=Level 5 (Sample #3)

Tests	Percent Agreement	Percent Disagreement	Chi Square
English			
BSM with CCT	66	34	.602
BSM with CERAS	66	34	.602
BSM with WNBD	70	30	2.389
CCT with CERAS	78	22	14.037 ***
CCT with WNBD	64	36	.194
CERAS with WNBD	71	29	4.843 *
Spanish			
BSM with CCT	70	30	5.918 *
BSM with CERAS	70	30	4.949 *
BSM with WNBD	71	29	6.792 **
CCT with CERAS	70	30	6.578 **
CCT with WNBD	71	29	8.381 **
CERAS with WNBD	76	24	14.648 ***

N = 75

* $p < .05$
** $p < .01$
*** $p < .001$

As indicated, more association in classifications was found in both English and Spanish analyses when Level 5 was used as the anchor for the Equivalent Percentile Cutoff Procedure. Whereas only two pairs of tests showed significant associations in English, all of the Spanish pairs were significantly associated. This may be due to the wider range of scores on all Spanish tests. Although there is more association in classification using Level 5 as the cutoff, between 22 and 36% of students are not consistently classified.

These significant associations may be misleading, however. High percent of agreement is expected because more than 75% of the students are classified as proficient. Moreover, the associations between measures should not overshadow that fact that between 22 and 36% of the students are inconsistently classified, even when the tests are all anchored at the same cutoff percentile.

Crosstabulations between bilingual classifications (RLP) based upon Level 5 as proficient were also performed. The results, listed in Table 33, indicate that only the CCT and CERAS were significantly associated. However, even where significant overlap in categorization occurs, 19% of the students are classified inconsistently. With the other test pairs as many as 26% of the students would be classified inconsistently.

Table 33

Results of Crosstabulations Between Bilingual Classifications (RLP) Proficient = Level 5 (Sample #3)

Tests	Percent Agreement	Percent Disagreement	Chi Square
BSM with CCT	77	23	.486
BSM with CERAS	79	21	.003
BSM with WNBD	79	21	.651
CCT with CERAS	81	19	9.050 **
CCT with WNBD	74	26	1.074

N = 75

**p < .01

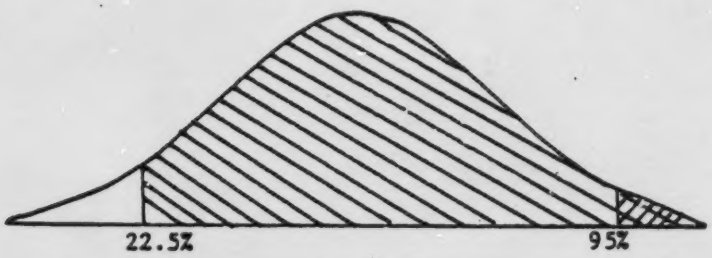
The apparently large numbers of students who would be inconsistently classified by any pair of tests anchored on the BSM Levels is surpassed only by the differences between classification if Level 5 or Level 6 are used. Tables 32 and 33 as well as Figures 1 and 2 illustrate these differences. If BSM II Level 6 is used as the cutoff for proficiency and the number of students scoring at this level used to determine the cutoff percentile, then 95% of the sample would be classified as limited English proficient; 5% would be proficient. On the other hand, if Level 5 is used, 22.5% would be classified as limited in English and 76.5% as proficient. For Spanish, using Level 6 as proficient, 67.5% would be classified as limited and 32.5% as proficient. Using Level 5, 26% would be limited and 74% proficient.

The English comparisons are most important because of the use of English language proficiency tests in school placement decisions. The difference between use of Level 5 or Level 6 would mean that 72.5% of

the students could be misclassified. This is about twice as devastating as if any two other language proficiency instruments had been used.

It is obvious that the use of the BSM Level designations is suspect. Consequently, most of the future analyses use total scores rather than BSM Levels.

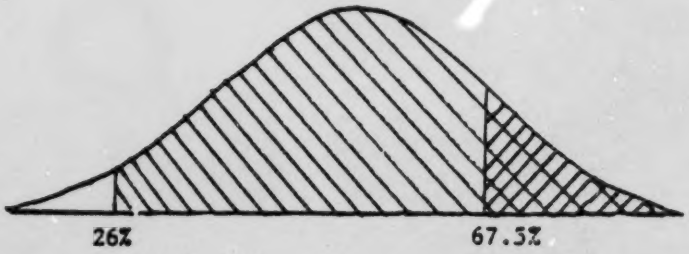
Figure 1: English CUTOFFS Using Level 5 and Level 6



= students classified as proficient in English when BSM II Level 5 used as cutoff

= students classified as proficient in English when BSM II Level 6 used as cutoff

Figure 2: Spanish CUTOFFS Using Level 5 and Level 6



= students classified as proficient in Spanish when BSM II Level 5 used as cutoff

= students classified as proficient in Spanish when BSM II Level 6 used as cutoff

In summary, the Equivalent Percentile Cutoff Procedure for equating language proficiency levels across tests was used to determine proficient vs. limited levels on the BSM, CCT, WNBD and CERAS. Crosstabulations of level designations between each pair of tests indicate that little association exists between tests and that approximately 1/4 to 1/3 of all students may be classified inconsistently using any pair of tests. Where association among measures was found, the results are theoretically predictable, in many cases. Measures of grammar and vocabulary are not strongly associated whereas communicative skills are related to both grammar and vocabulary.

The differences between use of Level 5 and Level 6 as designating proficiency in a language is suspect. Approximately 72% of students could be misclassified if Level 6 rather than Level 5 were used.

Language proficiency measures do not provide similar information about students; this is probably due to the non-unitary nature of "language proficiency" among children of this age/grade range. It may also be due to the use of highly suspect levels of the BSM as anchors for establishing cutoffs on all other tests and the inherent unreliability of individual measures.

Grade-Level Regression Procedures. Cutoff points based upon grade-level predictions were determined from individual regression analyses of each language test on achievement. Only students who had taken the CTBS (Districts 1, 2 and 4) were used to determine these cutoffs. There were two reasons for doing so: (1) 70-80% of the school districts in California use the CTBS as their yearly achievement battery; (2) one school district in this study used the Metropolitan Achievement Tests;

although separate analyses would have been appropriate for this sample, only 1 of 34 students reached grade level on the MAT; therefore, grade-level cutoffs could not be determined using this battery.

The score on each test which predicted grade level achievement was the cutoff point at which proficiency was determined. The cutoffs are listed in Table 34. Only English tests predicted English achievement. Insufficient sample size (N=10) of students with Spanish language proficiency and Spanish achievement data prohibited use of the grade-level regression procedure to establish Spanish levels.

Table 34

Grade-Level Cutoff Scores

Test	Proficient vs. Limited Cutoff Score	
	Proficient	Limited
BSM English Total Score	$\geq$ 19	$\leq$ 18
CCT English Total Score	$\geq$ 263	$\leq$ 262
CERAS English Total Score	$\geq$ 20	$\leq$ 19
WNBD English Total Score	$\geq$ 112	$\leq$ 111

After determining the English language proficiency cutoff levels, crosstabulations of student classifications for each pair of tests were performed. The results are indicated in Table 35.

Table 35

Grade Level Regression Procedure
Crosstabulations of Limited/Proficient Level Designations
(Sample #3)

Tests	Percent Agreement	Percent Disagreement	Chi Square
English only			
BSM with CCT	80	20	4.242 *
BSM with CERAS	75	25	.158
BSM with WNBD	89	11	.253
CCT with CERAS	69	31	.621
CCT with WNBD	80	20	3.659
CERAS with WNBD	74	26	.329

N = 75
*p < .05

Once again there is significant agreement between communicative skills and grammar; the Chi square for CCT with WNBD (vocabulary) has a probability of .06, approaching significance. While communicative competence and linguistic competence are somewhat related, even more so than discrete linguistic skills with each other, 20% of the students were classified differently by the pairs of tests. The importance of the association between these two types of measures should not, therefore, overshadow the 20 percent who are classified differently when a linguistic and a communicative test are used to measure language proficiency. This discrepancy may indicate that although the BSM and CCT share some common variance (measuring something in common) they also measure something very different. Hence 20% of the students are classified differently.

Correlations. Another procedure for examining the extent to which tests similarly classify students is through correlations among the tests. Correlation coefficients express the extent to which two tests may measure the same construct; the correlations between rankings will tell us to what extent students would be classified similarly.

The correlation matrix in Table 36 provides evidence to support the results of the above mentioned crosstabulations, of both the Equivalent Percentile and Grade-Level Regression cutoff procedures. Reliability of each instrument is provided (on the diagonals) to facilitate interpretation. As indicated, there is very little common variance (or association) among measures. Moreover, intercorrelations are substantially lower than the reliabilities of each individual instrument indicating that low correlations are not due to low test reliability. All of the tests appear to be measuring something different. The CCT and BSM (in English) are the only within language tests whose correlations did not reach significance. While the CERAS and BSM both appear to measure grammar, the former is receptive and the latter productive. WNBD measures vocabulary and CCT measures communication skills.

The highest overall correlations among English tests are with the CERAS. The CERAS is a receptive test; comprehension precedes production in language acquisition; because most of the students are in the process of acquiring English as a second language, it is logical that the test which best taps their current proficiency is the receptive one. The strongest measures in Spanish are both the BSM and WNBD. Because most of the students have Spanish as a first language, productive skills of grammar and vocabulary would be expected.

The BSM English does not function as well as the BSM Spanish; this is seen in the higher correlations between BSM Spanish and all other Spanish measures than between BSM English and other English measures. Highlighting this difference are the correlations between the CCT and BSM; for the Spanish forms, the correlation is .43 and significant.

The most obvious problem with the BSM English is its correlations with Spanish tests; the BSM English correlates higher with all Spanish language proficiency tests than it does with most of the English tests. This may indicate the lack of validity of the BSM English to measure underlying English language proficiency and more of a tendency to reflect L1 (Spanish) ability!

Table 36

Correlation Matrix of Language Proficiency Tests With
Internal Reliability Coefficients on the Diagonals
(Sample #3)

	CERAS-E	BSM-E	WNBD-E	CCT-E	CERAS-S	BSM-S	WNBD-S	CCT-S
CERAS-E	(.79)							
BSM-E[1]	.39***	(.74)						
WNBD-E	.24*	.31**	(xxx)					
CCT-E	.43***	.22	.26*	(.94)				
CERAS-S	.44***	.52***	.17	.16	(.79)			
BSM-S	.11	.70***	.39***	.12	.32**	(.89)		
WNBD-S	.13	.40***	.75***	.13	.31**	.55***	(xxx)	
CCT-S	.19	.33**	.17	.33**	.35**	.43***	.42*	(.90)

[1]BSM Total Scores are used

* $p < .05$

** $p < .01$

*** $p < .001$

There appear to be substantial method effects. Most tests correlated highest with their corresponding (other language) measure. This is true for CERAS-E, BSM-E, WNBD-E, BSM-S and WNBD-S. It may be that taking the test in one language teaches students how to take it in the other language, but this is unlikely, since such an effect would give rise to order effects, and none were found. The BSM (r = .70) and WNBD (r = .75) have the largest method effects; the CERAS (r = .44) and the CCT (r = .33) have the lowest. Tests with lower correlations between both forms may measure more underlying proficiency and less method.

The CCT has the lowest interlanguage correlation. Moreover, the correlation between both forms is lower than all correlations between the CCT Spanish and other Spanish tests and also lower than the correlation between the CCT English and CERAS English tests. This is further evidence that the CCT is measuring more underlying language proficiency and less method than the other measures.

The highest correlation between communicative competence (CCT) and linguistic competence (WNBD, BSM, CERAS) is .43. At the most, there is 43 percent common variance between linguistic and communicative competence, as measured by these instruments, for this population.

Because each test contributes unique information, measuring vocabulary, for example, will give a different picture of overall language proficiency than will measuring productive or receptive grammar or communication skills. One would expect that all measures of language proficiency measure something in common; they do. However, for children who are not yet literate, who are still learning their second language (and, perhaps, their first), language proficiency measures provide different pictures

of their skills. Each measure could classify a child differently.

Regression Analyses. Regression analyses, examining contributions of linguistic competence and communicative competence tests to prediction of achievement, were performed as indirect procedures for answering Question #3. If linguistic and communicative tests both contribute significantly to the prediction of achievement, further support could be lent to the previous findings—that different tests provide different information about language proficiency.

Stepwise Regression of achievement (CTBS English Total Battery) was performed on all English language proficiency measures (CCT, BSM, CERAS, WNBD). The Stepwise procedure (SAS, 1979) entered the variables into the regression equation in the order of decreasing contribution to total variance explained. Table 37 lists the results of this analysis.

Table 37

Stepwise Multiple Regression of CTBS/S, Level 1, English-Total Battery Achievement on Proficiency Tests

Step	Source	Simple R w/CTBS	ΔR^2	R^2	df	Overall F for regression	p <
1	CCT English	.48	.23	.23	1,25	7.44	.0115
2	WNBD English	.44	.09	.32	2,24	5.57	.0103
3	CERAS English	.38	.08	.40	3,23	5.17	.0007
4	BSM English	.40	.06	.46	4,22	4.74	.0065

N = 27 from Sample #3

The Stepwise regression analysis indicates that the CCT, the best predictor of achievement, explains 23 percent of the variance in achievement in this sample. An additional 23 percent was explained by the three

linguistic competence measures. In order to test whether given the CCT this additional contribution of the three linguistic competence measures was statistically significant, a test of the additional achievement variance accounted for by these measures combined was performed. This F ratio was computed from the sums of squares for regressions and for error, computed in the stepwise regression. A significant ($p < .05$) F ratio of 3.19 (df 3,22) was obtained. Given the CCT, the three linguistic measures significantly improved the prediction of achievement. This indicates that communicative competence and linguistic competence measures contribute unique information.

In order to test the reverse, that given the linguistic competence tests, the CCT would significantly improve achievement, a multiple regression was performed with the three linguistic measures entered first and then the CCT (Table 38). Results indicate that, given the three linguistic tests, the CCT did not significantly improve achievement. ($F = 3.45$, $df = 1,22$; $p < .077$).

The non-significant contribution of CCT to prediction of overall achievement may have been due to the inclusion of the math subtest of the CTBS in computation of the total battery score. Additional analyses were performed to determine if the contribution of the CCT would be significant using only language related achievement scores.

Table 38

Multiple Regression of CTBS English
Total Battery on English Language Proficiency
CCT Entered Last

Predictor	Source	Simple R w/CTBS	ΔR^2	R^2	df	Overall F for regression	p <
1	BSM English	.29	.09	.08	1,25	2.38	n.s.
2	CERAS English	.38	.09	.18	2,24	2.68	n.s.
3	WNBD English	.44	.20	.38	3,23	4.67	.0500
4	CCT English	.48	.08	.46	4,22	4.74	.0065

Two multiple regression analyses were performed with the dependent variable the CTBS Total Reading and Language Scores. In the first analysis, the three linguistic competence measures were entered first, followed by the CCT. Results of this analysis are listed in Table 39.

Table 39

Multiple Regression of CTBS English
Reading and Language Totals on English Language
Proficiency-Communicative Competence Test Entered Last

Predictor	Source	Simple R w/CTBS	ΔR^2	R^2	df	Overall F for regression	p <
1	BSM English	.26	.07	.07	1,26	1.84	n.s.
2	CERAS English	.44	.14	.21	2,25	3.27	n.s.
3	WNBD English	.46	.21	.42	3,24	5.79	.0100
4	CCT English	.55	.10	.52	4,23	6.32	.0014

Given the three linguistic competence measures the CCT does contribute significantly to the improvement of prediction of reading and language achievement. A significant ($p < .05$) F ratio of 5.02 with 1,23 degrees of freedom was found. This indicates that the CCT made a unique contribution and significantly improved prediction.

In the second analysis, the CCT was entered first and the three linguistic competence measures after it. Results are listed in Table 40.

Table 40

Multiple Regression of CTBS English Reading and Language Totals on English Language Proficiency-Linguistic Competence Tests Entered Last

Predictor	Source	Simple R w/CTBS	ΔR^2	R^2	df	Overall F for regression	p <
1	CCT English	.55	.30	.30	1,26	11.37	.0100
2	BSM English	.26	.06	.36	2,25	7.08	.0100
3	CERAS English	.44	.07	.43	3,24	5.93	.0100
4	WNBD English	.46	.09	.52	4,23	6.12	.0014

A significant ($p < .05$) ratio of 3.53 with 3,23 degrees of freedom was found. This indicates that, given the CCT, the three linguistic competence measures significantly improve the prediction of reading and language achievement.

Results of the regression analyses indicate that the CCT and the three linguistic competence measures provide unique information about language proficiency of students. Both types of information significantly improve prediction of achievement, especially in language and reading. Any one type of measure is not as good a predictor (or descriptor) as the

combination of measures. If students were classified using only one type of measure it is probable that many would be incorrectly classified.

Factor Analyses. Factor analysis is a statistical technique that has been used to investigate whether language proficiency is unitary or divisible into components that may be separately assessed (Oller & Perkins, 1979, 1980; Oller, 1980; Cummins, 1981b). Factor analytical procedures examine many correlations simultaneously to sort out patterns of relationships among variables. Through factor analysis we can examine whether the different language tests in this study all reflect the same underlying dimensions of language ability, or whether they tap different dimensions. Perhaps there is a general component common to all of the variances of the language tests but unique variances associated with each different test. The low to moderate intercorrelations among tests (Table 36) indicate that this may be the case.

Data from Sample #3 were used to investigate whether, for English and for Spanish, all tests load on one general factor. Included in the factor analyses are all language tests (BSM, CERAS, WNBD and CCT) as well as teacher judgments of students' communication skills and scores on the Cartoon Conservation Scales (CCS). The latter was included in the analyses in order to test Oller's (1978) hypothesized 'g' factor in language proficiency. Spanish and English tests were analyzed separately due to substantial interlanguage method factors as shown in Table 36. If a single factor analysis across languages were performed, the results would be complicated by method factors. With languages analyzed separately, method variance unique to each measure will not interfere with the analysis.

If, as Oller (1978) argues, one language-intelligence factor exists, we would expect a single factor with high factor loadings, or correlations of the individual tests with the "factor," for each variable.

Factor analysis using the Iterated Principal Axis Procedure (SAS, 1979) was performed. Two factors were extracted. Varimax rotation was then performed. Results of these procedures are indicated in Table 41.

Table 41

English Rotated Factor Pattern:
All English Language Tests, Teacher Ratings and CCS
Sample #3

Source	1st Factor	2nd Factor
BSM English	.61	.09
CERAS English	.71	-.02
WNBD English	.44	-.06
CCT English	.57	.10
Teacher Ratings	.28	.05
CCS	.05	.92

As indicated, the English language proficiency tests all have moderate factor loadings on a "language" factor. The CCS does not load on the language factor.

While these tests all measure something in common, communalities (Table 42) are substantially lower than estimated reliabilities (see diagonal entries, Table 36). This indicates that each test measures unique abilities distinct from those assessed by other tests, in addition to the one common factor.

Table 42

Communalities of Language Measures

English

Source	Communality Estimate R^2
BSM English	.38
CERAS English	.50
WNBD English	.20
CCT English	.33

Similar analyses for the Spanish measures were performed. Two factors were extracted (Table 43). Results confirm the findings of the English factor analysis. All Spanish language proficiency tests loaded on the first factor. Their communalities (Table 44) are substantially lower than each tests' estimated reliability (Table 36). Again, in addition to one common factor, each test measures unique abilities distinct from those assessed by other tests.

Table 43

Spanish Rotated Factor Pattern:

Spanish Language Tests and CCS

Sample #3

Source	1st Factor	2nd Factor
BSM Spanish	.78	.01
CERAS Spanish	.52	.11
WNBD Spanish	.74	.05
CCT Spanish	.58	.12
Teacher Ratings of Spanish CC	.12	.43
CCS	-.01	.82

N = 78

Table 44

Communalities of Language Measures

Spanish

Source	Communality Estimates R^2
BSM Spanish	.61
CERAS Spanish	.28
WNBD Spanish	.55
CCT Spanish	.35

The main difference between English and Spanish analyses is the loading of teacher ratings on the second factor for Spanish. This might indicate that teachers do not base Spanish language proficiency judgments on what they have heard. Rather, they use some non-verbal intelligence indicator and infer Spanish language proficiency from it. "This student is smart; he probably speaks Spanish well." Evidence for this interpretation is provided by comments made by bilingual teachers when filling out their rating lists. Many stated that they had difficulty rating the students' Spanish because they never heard them speak it.

The factor analyses of language tests indicate that while a common "language" factor appears to exist, each test contributes its unique variance. In practical terms, for this population, these tests provide different information about students. No two tests provide the same information. Interestingly, separate "linguistic competence" and "communicative competence" factors did not emerge. Perhaps the difference between these two types of "constructs" is exaggerated and somewhat artificial. It appears as though language proficiency is made up of many different types of skills. Separating these skills into two separate

constructs is not supported by these analyses.

Summary. The question of classification by two or more language proficiency instruments is beset with the problems inherent in determining level designations within each test. Attempting to approach the task in a variety of ways, five procedures for determining concordance among tests were performed. Two of these procedures, Equivalent Percentile Cutoffs and Grade-level Regression Cutoffs, use level designations to compare agreement and disagreement in classification by all pairs of tests. Due to the difficulty in determining what the levels "Proficient" and "Limited" mean, three procedures comparing all tests without use of levels were also performed. Correlations, Multiple Regressions and Factor Analyses were used to compare tests.

Results of all analyses indicate that for the population in question, no two language proficiency tests will give the same information about a child. That is, although some evidence of a common "language factor" exists, little common variance is found between measures. This indicates that the use of any one instrument for determining placement in special language programs may run the risk of substantial error in placement.

The measure of communication skills (CCT) was the best predictor of achievement (CTBS) and also correlated with measures of vocabulary (WNBD) and grammar (CERAS). While no two tests measure the same thing, the CCT appears to have more in common with vocabulary and grammar than they have with each other. Perhaps this indicates support for the belief that linguistic competence is part of communicative competence.

A final word of caution. The BSM Levels specified by the test authors are highly suspect. Because of the possibility of using Levels 5 or 6 as

the cutoff for proficiency, 72.5% of the students may be classified incorrectly. This could indicate that school districts wishing not to provide assistance to English as a second language learners could opt for the use of the lower level as indicating proficiency. Others, leaning in the opposite direction, could use Level 6. Because of the lack of discrimination among levels of the BSM II, the levels were not used for most of the analyses.

Question #4 What variables other than language proficiency differentiate between limited proficient bilingual and non-limited proficient bilingual students?

Limited proficient bilinguals (LPB) are students who appear to be limited in both L1 and L2. In order to determine what factors are concomitant with this dual language limitation, LPB students were compared with non-LPB students on a number of variables. These include both continuous variables (length of residency, parents' educational levels, achievement scores and cognitive development scores) and nominal variables (place of birth, language use with parents and siblings).

Students were grouped on the basis of RLP classifications according to both the BSM and CCT* (in both English and Spanish) using Level 6 as Proficient.** Equivalent Percentile Cutoff was used to determine proficient status on the CCT. Group 1 included those students (N=40) who were classified as LPB by both the BSM and CCT; Group 2 (N=27) included students

*The BSM and CCT act in very different ways, as evidenced in the analyses for Question #3. It was felt that because they are so different, students who were classified as LPB by both tests would be those students who were limited in verbal expression as well as productive grammar. Therefore, these two tests were selected.

**Analyses would have also been performed using Level 5 as the cutoff for proficiency. However, small sample size of LPB students (N=2) when BSM and CCT bilingual classifications were crosstabulated using Level 5 as the cutoff, prohibited these analyses.

who were classified as LPB by one instrument and not the other; Group 3 (N=13) was composed of students classified as proficient in either one or both language(s) as determined by agreement between the two tests. No PB students were identified through this crosstabulation (Table 45).

Table 45

Crosstabulations of Bilingual Classification by BSM II and CCT

BSM Bilingual Classification	CCT Bilingual Classification: LPB	Prof.Eng./ Lim. Span.	Prof. Sp/ Lim. Eng.	PB
LPB	40	1	9	2
Prof. Eng./ Lim. Span.	2	0	0	0
Prof. Span./ Lim. Eng.	13	0	10	1
PB	0	0	2	0

= Group 1

= Group 2

= Group 3

Comparisons among group means for the continuous variables using analyses of variance were performed; comparisons on nominal variables were made through t-tests. The ANOVAS (Tables 46-47) indicated that students' length of residency in the United States approached significance ($F=3.00$, $p < .06$). Perhaps with a larger sample of wider age-range students, length of residency would have been significant (Cummins, 1981).

Table 46

Analysis of Variance Results Comparing 3 RLP Groups on the CCS, CTBS English Total Battery, Student Length of Residency in the U.S. and Parents' Educational Levels

Source	Mean	S.D.	df	SS	Mean Square	F
CCS Total	20.82	5.78	2	.34	.17	.01
error				2469.11	33.37	
CTBS	163.48	50.88	2	5942.73	2971.37	1.15
error				67300.50	2588.48	
Yrs. in U.S.	7.32	2.21	2	29.43	14.72	3.00
error				367.55	4.90	
Yrs. Mother in School	5.82	3.04	2	9.55	4.77	.52
error				691.94	9.23	
Yrs. Father in School	5.88	3.46	2	18.78	9.39	.78
error				899.18	11.99	

Table 47

Means for Three RLP Groups

Variable	Group	N	Mean
CCS Total	1	37	20.81
	2	27	20.88
	3	13	20.69
CTBS Battery	1	9	143.33
	2	13	168.38
	3	7	180.29
Yrs. in U.S.	1	39	7.72
	2	27	6.48
	3	12	7.92
Yrs. Mother in school	1	39	5.97
	2	27	5.96
	3	12	5.00
Yrs. Father in school	1	39	5.82
	2	27	6.41
	3	12	4.92

No significant differences were found among groups on the variables of place of birth, language students use with parents, language parents use with students and language students use with siblings.

Students who appear to be limited in both L1 and L2 do not appear to be significantly different from bilingual students who are proficient in one or both languages. They may have lived in the United States less time. This is supported by Cummins' (1981) research on length of residency. In particular, students do not appear to suffer cognitively. Rather, all groups performed, on the average, above the exptected mean on the Cartoon Conservations Scales ($\overline{X}$ = 20.03). Academically, LPB students are different, although not significantly so, from their non-LPB peers. This lack of significance may have been attributable to small sample size.

However, when students who are limited in both languages are compared with those who are proficient in at least one language (Groups 1 and 3, respectively), a significant difference between groups was found on student place of birth (Table 48). Students proficient in English, Spanish or both were more frequently born in Mexico ($X^2 = 4.60$, $p < .05$).

Table 48

Crosstabulations of Groups 1 and 3
on Student Place of Birth

	Mexico	U.S.	
Group 1	15	23	LPB
Group 3	9	3	Proficient in at least one language

This is the only variable which significantly distinguished between the limited bilingual proficient students and their proficient peers.

Why should students' place of birth distinguish between these two groups of students?

In order to provide some insight into this relationship, students were regrouped by place of birth and t-tests performed to investigate differences in group means on the continuous variables; crosstabulations of these students on the nominal variables were also performed. Significant differences were found between groups (Tables 49-51).

Third grade Mexican born students have been in the U.S. for 6.5 years, on the average, compared to 8.4 years for their U.S. born bilingual peers. This difference is significant and confounded with place of birth. Students born in Mexico have significantly higher CCS scores than their

Mexican-American peers. Moreover, the scores of the Mexican students are higher on the CCT Spanish.

Table 49

T-Test Results of Comparisons of
Mexican and U.S. Born Bilingual Students

Variable	Group	N	Mean	S.D.	df	T
CCS Total	Mexico	34	23.08	4.57	71	3.19**
	U.S.	39	19.08	5.90		
Yrs. in U.S.	Mexico	35	6.51	2.02	73	-3.47***
	U.S.	40	8.20	2.16		
Comprehensive Test of Basic Skills (CTBS) Total Battery Read	Mexico	10	154.50	55.06	26	- .83
	U.S.	18	171.39	29.27		
CTBS Total Language & Reading	Mexico	10	86.36	43.12	28	- .97
	U.S.	19	101.39	49.27		
Yrs. Mother in School	Mexico	35	4.8	2.64	73	-2.90**
	U.S.	40	6.72	3.05		
Yrs. Father in School	Mexico	35	5.46	3.67	73	- .94
	U.S.	40	6.23			
CCT English	Mexico	34	227.76	32.60	69	-2.21*
	U.S.	37	242.62	23.57		
CCT Spanish	Mexico	32	253.28	28.18	69	2.22*
	U.S.	38	236.71	33.27		

* $p < .05$
** $p < .01$
*** $p < .001$

Table 50

Crosstabulations of Student Place of Birth
With Student Language Use With Parents

Student Place of Birth	Language Student Uses With Parents: English	Spanish	Both
Mexico	0	27	8
U.S.	3	19	18

Table 51

Crosstabulations of Student Place of Birth
With Parent Language Use With Student

Student Place of Birth	Language Parents Use With Student: English	Spanish	Both
Mexico	0	28	2
U.S.	4	24	6

In both crosstabulations, results indicate that students born in Mexico do not use English as the only language of interaction with their parents. Chi square analysis of differences between groups on these language use variables indicate that students born in Mexico use significantly less English with their parents ($X^2 = 6.92$ with 1 df, $p < .01$); their parents use significantly less English with the students

($\chi^2 = 5.41$ with 1 df, $p < .02$)*

Mexican born students use significantly less English in interactions with their parents. Perhaps this language use at home is related to the significant difference, in favor of the Mexican born students, in scores on the CCT and CCS.

In order to investigate the relationship between home language use and the variables of language proficiency and cognitive functioning, t-tests were performed which compared students and parents (those who use only Spanish at home with those who report using both Spanish and English) with respect to language proficiency and cognitive functioning. As indicated in Tables 52-54, no relationship between student or parent language use at home and cognitive functioning was found. Nor did language use at home have any strong and consistent relationship with language proficiency.

The relationship between place of birth, CCS scores and the CCT scores (Table 49) does not appear to be attributable to home language use. Perhaps another interpretation is possible. The CCS measures cognitive development. If the CCT also picks up developmental influences, then age of student may explain this relationship. Perhaps Mexican born students are retained in school more frequently or enter school at later ages. They may then be more cognitively developed than their United States born peers.

* Decrease in sample size, though uniform for both groups, was due to an error in data collection. The parent questionnaire item used to elicit the information of parent language use was free response; the item on student language used was forced choice. Differences in sample size notwithstanding, results are consistent.

Table 52

T-test Comparisons of Language Proficiency and Cognitive Functioning Scores for Students Who Use Spanish Only (S) vs. Spanish and English (S & E) at Home

Variable	Group	N	Mean	S.D.	df	T
CCS Total	S	46	20.50	5.81	71	- .9492
	S & E	27	21.81	5.53		
CCT English	S	44	232.07	31.66	69	-1.4174
	S & E	27	242.15	24.25		
CCT Spanish	S	43	249.93	31.38	68	1.1235
	S & E	27	241.78	26.34		
WNBD English	S	46	63.48	29.34	71	.4100
	S & E	27	60.78	22.93		
WNBD Spanish	S	45	62.93	25.34	69	2.2428 *
	S & E	26	49.65	21.56		
CERAS English	S	46	17.59	2.47	71	-1.0272
	S & E	27	18.15	1.81		
CERAS Spanish	S	46	17.96	2.29	71	.7568
	S & E	27	17.52	2.55		
BSM English	S	46	13.39	3.88	71	.8694
	S & E	27	12.63	3.09		
BSM Spanish	S	46	13.85	4.45	71	2.5369 *
	S & E	27	11.04	4.77		

* $p < .05$

Table 53

T-test Comparisons of Student Language Proficiency and Cognitive Functioning Scores for Parents Who Use Spanish Only (S) vs. Spanish and English (S & E) at Home

Variable	Group	N	Mean	S.D.	df	T
CCS Total	S	52	21.17	5.69	58	- .3944
	S & E	8	22.00	4.12		
CCT English	S	50	231.06	29.92	56	-1.9859
	S & E	8	252.63	15.49		
CCT Spanish	S	49	247.61	28.10	55	.9564
	S & E	8	236.50	43.36		
WNBD English	S	52	58.38	22.63	58	-1.8909
	S & E	8	75.63	32.32		
WNBD Spanish	S	50	56.28	23.91	56	.4706
	S & E	8	52.00	23.65		
CERAS English	S	52	17.63	2.17	58	-1.3992
	S & E	8	18.75	1.49		
CERAS Spanish	S	52	18.04	2.40	58	2.0314 *
	S & E	8	16.13	3.00		
BSM English	S	52	13.02	3.39	58	1.8879
	S & E	8	10.38	4.34		
BSM Spanish	S	52	13.00	4.75	58	1.1815
	S & E		10.86	4.61		

* $p < .05$

Table 54

T-test Comparisons of Student Language Proficiency and Cognitive Functioning Scores for Combined Student and Parent Home Language Use for Students Who Use Same Language(s) as Parents

Variable	Group	N	Mean	S.D.	df	T
CCS Total	S	38	21.00	5.72	43	- .4998
	S & E	7	22.14	4.41		
CCT English	S	36	228.67	31.06	41	-1.8540
	S & E	7	251.14	16.11		
CCT Spanish	S	35	247.74	31.60	40	- .2155
	S & E	7	250.43	19.56		
WNBD English	S	38	58.50	23.72	43	-1.1053
	S & E	7	69.71	29.88		
WNBD Spanish	S	37	59.89	24.14	42	.4786
	S & E	7	55.14	23.67		
CERAS English	S	38	17.66	2.30	43	-1.3150
	S & E	7	18.86	1.57		
CERAS Spanish	S	38	17.92	2.41	43	1.0779
	S & E	7	16.86	2.34		
BSM English	S	38	13.24	3.93	43	1.1391
	S & E	7	11.43	3.40		
BSM Spanish	S	38	13.66	4.58	43	1.4678
	S & E	7	10.86	4.98		

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

Overview

The phenomenon of limited bilingual language proficiency, or "semi-lingualism" as it is referred to in the European literature, is reported to have negative academic and/or cognitive consequences (Cummins, 1978; Skutnabb-Kangas & Toukomaa, 1976; Toukomaa & Skutnabb-Kangas, 1977; Duncan & DeAvila, 1978a). Among Mexican-American minority language students in the United States, the phenomenon has been cited as widespread by some (Dulay & Burt, 1980) and negligible by others (Spiegel & Foster, 1980). Where such students have been diagnosed, fears of concomitant cognitive retardation are also voiced. These students are believed to be in need of special education. The belief is that cognitive abilities rather than language abilities are in need of remediation.

In order to better understand the phenomenon of "limited bilingual language proficiency" and its possible cognitive/academic consequences, the following research questions were proposed:

1. Among students classified as "limited proficient bilingual," is the total repertoire of grammatical skills distributed across languages? If so, is this distribution complementary?

2. Among students classified as "limited proficient bilingual," is the total repertoire of vocabulary and communication skills distributed across languages as a function of domain?

3. Among students classified as "limited proficient bilingual," is there a difference in assigned bilingual proficiencies when different types of proficiency tests (linguistic competence vs. communicative competence) are used to assess verbal behavior?

4. What variables other than language proficiency (i.e., cognitive development, achievement, length of residency) differentiate between limited proficient bilingual and non-limited proficient bilingual students?

The results of this investigation, which studied third grade Mexican-American students from San José, California, the area in which Dulay and Burt (1980) report the highest occurrence of limited bilingual proficient students, do not lend support for the conclusion that limited bilingual language proficiency is a widespread phenomenon. Rather, the conclusions implicate the use of language tests which measure only one aspect of language (i.e., grammar or vocabulary) and which provide cutoffs (levels) of questionable validity. Results also indicate that, among students who do exhibit this dual language limitation, no cognitive deficits are found.

Conclusions

No single language proficiency measure accounts for the various language skills that students possess. Results of crosstabulations between test pairs, intercorrelations among language tests, multiple regressions and factor analyses indicate that each language test used in this investigation provides unique information about the language proficiency of students. Measures of grammar, for example, do not indicate the total language proficiency of students.

Within the oral/aural aspects of language, no one "global language factor" accounts for differences measured by all of the tests. Communicative competence, as measured by the CCT, and linguistic competence, as measured by the CERAS, BSM and WNBD tasks, while sharing some common attributes, make unique contributions to the measurement of language proficiency. The moderate relationships between the CCT and vocabulary

or between the CCT and grammar indicate that communicative competence includes these discrete linguistic skills. However, it also contributes its own unique information. This result has been found by other researchers (Overall, 1978; Politzer et al., 1981; Cummins, 1981b). While Oller argues for one "global language factor" underlying proficiency, he also indicates that oral speaking proficiency may contribute unique variance (Oller & Hinofotis, 1980).

Complicating the issue of various tests measuring different skills is the lack of comparability in the level designations provided by test developers. The cutoff points used to designate limited and proficient status in a language are not anchored on criteria which indicate whether students can or cannot function effectively in English-only instruction. As currently available, many language proficiency instruments fail to distinguish between those who may be truly limited or proficient in a language and those who are not. This problem could be resolved if specific external criteria (i.e., grade level achievement) were used to validate cutoff scores and level designations. Any evaluative tool should be valid for the purposes for which it is used. Language proficiency tests which designate proficiency levels based on monolingual norms on the test are not necessarily valid indicators of language proficiency sufficient to succeed in English-only classrooms, without special language assistance.

These two problems, the concurrent and content validity of language tests and the predictive validity of cutoff scores, lead to the possible misclassification of non-English language background (NELB) students. Approximately thirty percent of the students in this study were inconsistently classified using any two instruments. This same result was

found by Ulibarri, Spencer and Rivas (1981), who compared English oral language tests and classifications of NELB students. This investigation compared both English and Spanish tests and found similar results.

Misclassification of immigrant students is a predictable occurrence, given current test norming procedures and the sociolinguistic variables of the testing situation (Labov, 1970). Abudarham (1980) provides evidence for this phenomenon in citing a report out of England which reviewed procedures and problems associated with assessment of West Indian students: "Where children are suspected as being wrongly placed. . . this is four times as likely in the case of immigrant students." Vernon (1957) reports that "the rate of misclassification in the selection of children for secondary schools, even in normal children, was as high as 10% and that the probability of error in handicapped children and 'difficult to test population' is considerably higher."

Error in classification might occur if a bilingual's language skills are measured without consideration of distribution of skills across two languages. It was thought that perhaps students in San José, California had been raised in diglossic environments using Spanish for one set of functions and English for another. Results of this investigation indicate that these students do not appear to come from diglossic bilingual environments. Rather, students use English and Spanish in both home and school situations; although stronger in Spanish, their languages do not seem to be separated as a function of domain. This result has been supported by another researcher who concluded that the bilingualism of the Mexican-American population of San José is generally characterized by an absence of diglossia (Beltramo, 1972). The strength

of this conclusion, however, must be tempered by the fact that home language was measured in a curious, perhaps less than ideal manner. Responding to pictures and puppet interactions of home situations, students were asked to use language within this domain. What may have been measured on these tasks was students' school use of home language rather than their true home language use. Measurement of language use in the home might provide a more valid indication of whether diglossia indeed exists among these students. Moreover, diglossia and the possible complementary distribution of language skills might be more accurately assessed through comparisons of actual elicited vocabulary, rather than through word counts.

The lack of distribution of language skills across languages and domains might also be related to age and length of residency of students. Due to the restricted range of the third grade, there was little variation in length of residency.

An important conclusion to be drawn from comparing these students' vocabulary by domain is the possible need for inclusion of the home domain in language tests which measure vocabulary. In this investigation, students' home vocabulary appeared to be stronger than their school vocabulary. If measurement only includes school-related vocabulary, students may appear limited because only part of their vocabulary knowledge was tapped.

Finally, the following conclusions related to the concept of "semilingualism" seem warranted: (1) By definition, "semilingualism" is limited proficiency in both L1 and L2. The results of this investigation indicate that "semilingualism" exists only as a concept. There is no reliable evidence for the existence of the phenomenon.

Since there appears to be no way to reliably measure the phenomenon nor unambiguously identify it, it is not useful as a construct. Observations used to indicate the existence of "semilingualism" may be the result of limited language proficiency instruments rather than limited language proficiency. (2) Some of the literature on "semilingualism" suggests that limited bilingual language proficiency may have negative cognitive results. No indication of this relationship was found to exist for this population. Although any one test or combination of tests indicated that a student may appear to be limited in both L1 and L2 (and this identification was likely to be unreliable and of questionable validity), no differences in cognitive development were discovered between limited and proficient groups of bilingual students.

Why are these results different from those found in the literature? The most obvious conclusion lies in the difference in operationalization of the constructs "language proficiency" and "cognitive ability". In the Finnish research, the two seem to be confused. When Finnish immigrant students living in Sweden were compared on language proficiency, the predominant measure of language was the verbal ability section of the Finnish version of the Thurstone Primary Mental Abilities test series (Skutnabb-Kangas & Toukomaa, 1976). In this research, the distinction between cognitive ability and language proficiency cannot be drawn.

In this investigation, the CCS was used as a measure of cognitive ability; the test is unaffected by language backgrounds of the students. It is not susceptible to bias due to language. In most of the literature, however, when reference to the relationship between cognitive

functioning and language proficiency is made, cognitive functioning, it can be inferred, means verbal I.Q. As Strump (1978) has stated: "Indeed, it has recently been empirically demonstrated that variance in language proficiency accounts for the main portion of variance in tests of intelligence."

Perhaps the "lower cognitive functioning" attributed to immigrant or NELB students who "fail to develop proficiency in L1 and L2" may actually be an inappropriate conclusion based on the results of verbally loaded measures of cognitive functioning and language proficiency measures which may lack predictive and/or content validity. It is obvious and logical that verbally loaded cognitive tests will show a strong relationship with other language related tasks.

Limited bilingual language proficiency (or semilingualism) is a construct which lacks sufficient evidence to be accepted as valid. No single language proficiency measure accounts for the various language skills that students possess. That is, students appear to be "semi-lingual" as a function of the type of test chosen and the cutoff points used to designate limited or proficient status in the language(s). If the cutoffs are arbitrary, they cannot be used as indicators to validate a construct. Limited bilingual language proficiency may be an artifact induced by limited bilingual measurement instruments.

Recommendations

The use of language proficiency tests for placement of non-English language background students in the United States is a widespread and, perhaps, overdeveloped occurrence. Consequences of misclassification due to an inadequate or inappropriate test may include improper

placement, insufficient instruction, and, as a result, lower academic achievement than would have been reached if appropriate instruments and proper placement had occurred. The ultimate consequence of such testing may include increased drop-out rates and perpetuation of unequal educational opportunity for language minority students. This is the opposite of the result which the use of language proficiency tests for placement decisions is intended to have.

In order to appropriately place and educate non-English language background students, accurate assessment is essential. The following recommendations, which are based upon the conclusions of this research, call for further research in language tests, their predictive validity, and reclassification criteria for bilingual students, as well as the use of multiple indicators of language proficiency for placement decisions. These recommendations are similar to those made by DeAvila, Carvantes and Duncan (1978), DeAvila et al. (1978), Ulibarri, Spencer and Rivas (1981), and NIE (1981):

(1) Recommendations for replication of this investigation include the following: (a) students from various locales, specifically those closer to the Mexican border or other ports of entry, should be studied; (b) age data should be collected; for immigrant students who enter school late or who are retained in school, grade level does not necessarily indicate age; (c) parent occupation and length of residency data may help in predicting both L1 and L2 proficiency of students.

(2) School placement decisions should be made only after evaluation of linguistic and communicative aspects of language proficiency has been performed. This can be achieved through the use of multiple measures which include grammar, vocabulary and communication skills or

through the use of a single measure of verbal expression which incorporates these areas.

(3) Levels of language proficiency which are used as cutoffs for determining limited or proficient status in language(s) should not be accepted as valid simply because the test authors provide them for use. Rather, more valid level indicators, such as grade-level prediction, should be developed and used as criteria for cutoffs.

(4) Some tests, such as the BSM or CCT, are scored on the basis of subjective judgments. Although the BSM purports to be objective in its scoring (Hernández-Chávez, 1978), moderate interscorer reliabilities indicate that scoring of this test is not highly reliable. The CCT is rated subjectively. Both types of tests, if used for placement decisions, lack sufficient reliabilities to consistently classify individual students. They can be used reliably for group decisions, but decisions about individual students should be made in conjunction with other indicators of language proficiency. Extensive rater training is necessary to make individual placement decisions using either the BSM or the CCT. Similarly, some tests are more sensitive to test administrators. Tests which are affected by tester should be administered carefully and their results used with caution.

(5) The different language skills found on the linguistic and communicative measures may actually reflect differences due to test format or scoring system (integrative vs. discrete point scoring). Further research such as the multitrait/multimethod work of Bachman and Palmer (1981) may be necessary to sort out method from trait. For the purposes of testing for placement, however, with appropriate interrater reliabilities, the use of multiple tests measuring listening,

speaking, grammar, vocabulary and communication skills (rated globally or scored discretely) will provide a general picture of students' language proficiency.

(6) Only one factor of language proficiency was found. This might indicate that communicative competence and linguistic competence are not separate constructs. As measured, however, linguistic skills and communicative skills provide different information about students' language proficiency. Research efforts to validate the construct of communicative competence may be warranted; such work should include definitions of communicative competence and should operationalize these definitions through a variety of communicative tasks.

(7) Cummins (1981) favors the use of context-reduced cognitive-academic-language-proficiency (CALP) type measures over the traditional context-embedded basic-interpersonal-communication-skills (BICS) type language measures for placement and reclassification of NELB students. Perhaps these CALP type measures would be those like the CERAS, which taps, besides receptive control of grammar, analogical reasoning and test-taking skills. The results of this research indicate, contrary to Cummins' hypothesis, that, for this age student, an apparently more BICS type measure, the CCT, was the best predictor of academic achievement. Perhaps, while this test appears to measure BICS, it taps CALP through the inclusion of school related tasks. More research is needed to disentangle these two theoretical constructs, develop tools to measure them, and then to validate the hypothesized relationships with achievement.

(8) This investigation measured oral/aural language proficiency skills. Perhaps results would be different if written skills (reading

and writing) were also measured and included in the language tests. For preliterate students, however, the use of oral/aural indicators may be more practical.

(9) Results of this study should only be generalized to the population sampled; limited bilingual language proficiency could be a phenomenon which occurs with older or younger students. More research is needed to understand the relationship between age of student, length of residency, home language use, language development and achievement. Results of research by Politzer et al. (1981) suggest that linguistic competence and communicative competence tests may differentially predict achievement in older or younger students. These results may be attributable to the acquisition versus learning distinction made by Krashen (1981). Research which attempts to validate this distinction and its relation to L2 proficiency seems warranted.

(10) Specific research to investigate home language use and language development among bilingual students who differ in length of residency, place of birth and cognitive development may provide insights into practices related to successful language development. This information could be used in parent education as well as other early intervention programs.

BIBLIOGRAPHY

Abudarham, S. "The Role of the Speech Therapist in the Assessment of Language-learning Potential and Proficiency of Children with Dual Language (DL) Systems or Backgrounds." Journal of Multi-Lingual and Multicultural Development 3 (1980): 187-206.

Albin, A. and Alexander, R. The Speech of Yugoslav Immigrants in San Pedro, California. The Hague: Nijhoff, 1972.

Allen, J. and Davies, A. eds. Testing and Experimental Methods. The Edinburough Course in Applied Linguistics, Vol. #4. London: Oxford University Press, 1977.

Allen, R.R. and Brown, K.L. eds. Developing Communication Competence in Children. Skokie, Ill.: National Textbook Co., 1976.

Anastasi, A. Psychological Testing, 4th Edition. New York: Macmillan, 1976.

Anisfeld, E. A Comparison of the Cognitive Functioning of Monolinguals and Bilinguals. Ph.D. dissertation, McGill University, 1964.

Bachman, L. and Palmer, A. The Construct Validation of Tests of Communicative Competence. Paper prepared for Language Proficiency Assessment Symposium organized by InterAmerica Research Associates, Inc., Airlie House, Virginia, March 14-18, 1981.

Balkan, L. Les effets du bilinguisme francais-englais sur les aptitudes intellectuelles. Bruxelles: Aimav, 1970.

Barker, G.C. Social Functions of Language in a Mexican-American Community. Ph.D. dissertation, University of Chicago, 1947.

Bartz, W.H. A Study of the Relation of Certain Learner Factors with the Ability to Communicate in a Second Language (German) for the Development of Measures of Communicative Competence. Ph.D. dissertation, Ohio State University, Columbus, 1974.

Bay Area Bilingual Education League (BABEL). Bilingual Testing and Assessment: Proceedings of BABEL Workshop and Preliminary Findings Multilingual Assessment Program. Berkeley, Ca.: BABEL, 1972.

_____. Concurrent Validity Study: BSM and BOLT. Berkeley, Ca.: BABEL, 1979.

Beltramo, A.F. Lexical and Morphological Aspects of Linguistic Acculturation by Mexican-Americans in San Jose, California. Ph.D. dissertation, Stanford University, 1972.

_____ and de Porcel, A. "Some Lexical Characteristics of San Jose Spanish." In El Lenguaje de los Chicanos, pp. 122-137. Edited by Eduardo Hernández-Chávez, Andrew Cohen and Anthony Beltramo. Arlington, Va.: Center for Applied Linguistics, 1975.

Bennett, A. and Slaughter, H. A Sociolinguistic/Discourse Approach to the Description of the Communicative Competence of Linguistic Minority Children. Paper prepared for Language Proficiency Assessment Symposium organized by InterAmerica Research Associates, Inc., Airlie House, Virginia, March 14-18, 1981.

Ben-Zeev, S. The Influence of Bilingualism on Cognitive Development and Cognitive Strategy. Ph.D. dissertation, University of Chicago, 1972.

Blom, J. and Gumperz, J. "Social Meaning in Linguistic Structures: Code-Switching in Norway." In Directions in Sociolinguistics: The Ethnography of Communication, pp. 407-434. Edited by John J. Gumperz and Dell Hymes. New York: Holt, Rinehart & Winston, 1972.

Bloomfield, L. Language. New York: Holt, Rinehart & Winston, 1933.

Board of Education of the City of New York. Language Assessment Battery (LAB). New York: Houghton-Mifflin, 1976.

Bowen, D.J. "Local Standards and Spanish in the Southwest." In Studies in Language and Linguistics, pp. 154-163. Edited by Ralph E. Ewton, Jr. and Jacob Ornstein. El Paso, Texas: Western Press, 1972.

Brent-Palmer, C. "A Sociolinguistic Assessment of the Notion 'Immigrant Semilingualism' from a Social Conflict Perspective." Working Papers in Bilingualism, No. 17, April 1979 : 135-180.

Briere, E.J. "Are We Really Measuring Proficiency With Our Foreign Language Tests?" Foreign Language Annals 4 (1971): 385-391.

Brown, M.E. and Zirkel, P.A. "Emerging Instrumentation for Assessing Large Dominance." In Occasional Papers on Linguistics: Proceedings of the First International Conference on Frontiers in Language Proficiency and Dominance Testing, pp. 202-225. Edited by James E. Redden. Southern Illinois University at Carbondale, 1977.

Burt, M. and Dulay, H. "Some Guidelines for the Assessment of Oral Language Proficiency and Dominance." TESOL Quarterly 12 (1978):177-192.

_____. "Hispanic Students May be Misdiagnosed." California Teacher. California Federation of Teachers, April 1980.

_____ and Hernández-Chávez, E. Bilingual Syntax Measure: Technical Handbook. English edition. New York: Harcourt, Brace, Jovanovich, Inc., 1975.

_____ and Taleporos, E. Bilingual Syntax Measure II: Technical Handbook. New York: Harcourt, Brace, Jovanovich, 1980.

Bye, T. Tests That Measure Language Ability. Berkeley, Ca.: BABEL/LAU Center, 1978.

California State Department of Education. Bilingual Program, Policy and Assessment Issues. Sacramento, Ca.: State Department of Education, 1980.

California State Statutes. Bilingual Education Improvement and Reform Act (Chapter 1339), 1980.

Canale, M. A Communicative Approach to Language Proficiency Assessment in a Minority Setting. Paper prepared for Language Proficiency Assessment Symposium organized by InterAmerica Research Associates, Inc., Airlie House, Virginia, March 14-18, 1981.

_____ and Swain, M. "Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing." Applied Linguistics 1 (1980): 1-47.

Cardenas, D.N. "How Proficient is the Bilingual Student in His Bilinguality?" In Teaching Spanish to the Spanish Speaking: Theory and Practice, pp. 191-198. Edited by Guadalupe Valdes Fallis and Rodolfo Garcia-Moya. San Antonio, Texas: Trinity University, 1976.

Carroll, J.B. "Research on Teaching Foreign Languages." In Handbook of Research on Teaching, pp. 1060-1100. Edited by N.L. Gage. Chicago: Rand McNally, 1963.

_____. "Foreign Language Proficiency Levels Attained by Language Majors Near Graduation From College." Foreign Language Annals, 1 (1967): 131-151.

_____. "Fundamental Considerations in Testing for English Language Proficiency of Foreign Students." In Teaching English as a Second Language, 2nd edition, pp. 313-320. Edited by H.B. Allen and R.N. Campbell. New York: McGraw-Hill, 1971.

Carter, T.P. Mexican-Americans in School: A History of Educational Neglect. New York: College Entrance Examination Board, 1970.

Cazden, C. "The Neglected Situation in Child Language Research and Education." In Language and Poverty, pp. 81-101. Edited by F. Williams. Chicago: Markham, 1970.

_____; John, V.P. and Hymes, D. eds. Functions of Language in the Classroom. Columbia University, N.Y.: Teachers College Press, 1972.

Center for Bilingual Education, Northwest Regional Educational Laboratory. Assessment Instruments in Bilingual Education. National Dissemination and Assessment Center, California State, Los Angeles, 1978.

Cervantes, R. and Nakano, P. Oral Language Tests: The Language Assessment Scales and Bilingual Syntax Measure. Unpublished paper, 1979.

Chesarek, S. Cognitive Consequences of Home or School Education in a Limited Second Language: A Case Study in the Crow Indian Bilingual Community. Paper prepared for Language Proficiency Assessment Symposium organized by InterAmerica Research Associates, Inc., Airlie House, Virginia, March 14-18, 1981.

Christian, D. and Tripp, R. "Teachers' Perceptions and Children's Language Use." In Final Report to Carnegie Corporation. Children's Functional Language and Education in the Early Years. Arlington, Va.: Center for Applied Linguistics, 1978.

Clark, J. Foreign Language Testing: Theory and Practice. Philadelphia: Center for Curriculum Development, 1972.

_____. "Theoretical and Technical Considerations in Oral Proficiency Testing." In Testing Language Proficiency, pp. 10-28. Edited by R. Jones and B. Spolsky. Arlington, VA.: Center for Applied Linguistics, 1975.

Cohen, Andrew D. A Sociolinguistic Approach to Bilingual Education. Rowley, Mass.: Newbury House Publishers, 1975.

_____ and Roll, C.L. Assessing Bilingual Speaking Skills: In Search of Natural Language. No. 6. Papers in Applied Linguistics, Evaluating Evaluation. Arlington, Va.: Center for Applied Linguistics, 1979.

Cooper, R. "Two Contextualized Measures of Degree of Bilingualism." Modern Language Journal 53 (1969): 172-178.

Cox, B. Working manuscript. Stanford University, 1980.

Cronbach, L.J. Essentials of Psychological Testing, 3rd edition. New York: Harper & Row, 1970.

CTB/McGraw Hill. Comprehensive Tests of Basic Skills, Level 1, Form S. Monterey, Ca.: CTB/McGraw Hill, 1974.

_____. Comprehensive Tests of Basic Skills, Level 1, Espanol. Monterey, Ca.: CTB/McGraw Hill, 1979.

Cummins, J. "Educational Implications of Mother Tongue Maintenance in Minority Language Children." Canadian Modern Language Review, 34 (1978): 395-416.

_____. "Linguistic Interdependence and the Educational Development of Bilingual Children." Review of Educational Research 49 (1979a): 222-251.

_____. "Cognitive/Academic Language Proficiency, Linguistic Interdependence, The Optimum Age Question and Some Other Matters." Working Papers in Bilingualism, No. 19 (Nov. 1979b): 197-205.

_____. "The Entry and Exit Fallacy in Bilingual Education." Paper presented at the Ninth Annual NABE Conference, Anaheim, Ca., April 1980.

_____. "Age on Arrival and Immigrant Second Language Learning in Canada: A Reassessment." Applied Linguistics 2 (1981a): 132-149.

_____. "Linguistic Interdependence Among Japanese Immigrant Students." Paper presented at Language Proficiency Assessment Symposium organized by InterAmerica Research Associates, Inc., Airlie House, Virginia, March 14-18, 1981b.

_____. "The Role of Primary Language Development in Promoting Educational Success for Language Minority Students." Paper presented for the California State Department of Education Compendium on Bilingual-Bicultural Education, 1981c.

_____ and Gulutsan, M. Some Effects of Bilingualism on Cognitive Functioning. Mimeo, Edmonton, Alberta, Canada: University of Alberta, 1973.

Cziko, G. "Psychometric and Edumetric Approaches to Language Testing: Implications and Applications." Applied Linguistics 2 (1981): 27-44.

Dailey, J.T. Language Facility Test. Alexandria, Va.: The Allington Corporation, 1965.

Dale, P.S. Language Development: Structure and Function, 2nd edition. New York: Holt, Rinehart & Winston, 1976.

Darcy, N. "The Effect of Bilingualism Upon the Measurement of the Intelligence of Children of Preschool Age." Journal of Educational Psychology 37 (1946): 21-44.

Davies, A. "Language Testing (Survey Article)." Language Teaching and Linguistics: Abstracts 2 (1978): 145-159, 215-231.

DeAvila, E. Cartoon Conservation Scales: Administration Manual. Corte Madera, Ca.: Linguametrics Group, 1980a.

_____. Cartoon Conservation Scales: Technical Manual. Corte Madera, Ca.: Linguametrics Group, 1980b.

_____. "Relative Language Proficiency Types: A Comparison of Prevalence, Achievement Level and Socio-Economic Status." Report submitted to Rand Corporation, February 1981.

_____ et al., "Predicting the Academic Success of Language Minority Students from Developmental, Cognitive Style, Linguistic and Teacher Perception Measures." Paper reporting CUTY supported by NIE Contract #400-65-0051, Southwest Educational Development Laboratory, June 1979.

_____; Cervantes, R.A., and Duncan, S.E. "Bilingual Program Exit Criteria." Paper submitted to the Office of Program Evaluation and Research, California State Department of Education, May 1978.

_____ and Duncan, S.E. Language Assessment Scales--LAS I, 2nd edition. Corte Madera, Ca.: Linguametrics Group, 1977.

_____. Definition and Measurement: The East and West of Bilingualism. Corte Madera, Ca.: Linguametrics Group, 1978a.

_____. "A Few Thoughts About Language Assessment: The LAU Decision Reconsidered." Bilingual Education Paper Series: National Dissemination and Assessment Center, California State, Los Angeles, 1978b.

_____. "Bilingualism and the Metaset." NABE Journal 3 (1979): 1-20.

_____ and Havassy, B. "The Testing of Minority Children - A Neo-Piagetian Approach." In Bilingual Education, pp. 381-384. Edited by H. La Fontain, et al. Wayne, N.J.: Avery Publishing Group, Inc., 1978.

_____ and Pulos, S. "Bilingualism and Cognitive Development." In Piagetian Theory and Its Implications for the Helping Professions. pp. 124-139. Edited by M. Poulsen & G. Lubin. Los Angeles: University of California, 1979

_____ and Ulibarri, D. Letter to the Editor, NABE News 4 (1981): 2ff.

Department of Health, Education and Welfare (DHEW). Request for Proposal: Development of Entry/Exit Criteria and Associated Assessment Procedures for Bilingual Education Projects (RFP Number: 78-111), August 1978.

Dieterich, T.; Freeman, C. and Crandall, J. "A Linguistic Analysis of Some English Proficiency Tests." TESOL Quarterly 13 (1979): 535-550.

Dissemination Center for Bilingual Bicultural Education (DCBBE). Evaluation Instruments for Bilingual Education: An Annotated Bibliography. Austin, Texas, 1975.

Dulay, H. and Burt, M. "The Relative Proficiency of Limited English Proficient Students." In Current Issues in Bilingual Education, pp. 181-200. Edited by J.E. Alatis. Georgetown University Round Table on Languages and Linguistics, 1980. Washington, D.C.: Georgetown University Press, 1980.

_____. "Aspects of Bilingual Education for LES/NES Students." Bilingual Program, Policy and Assessment Issues, CSDE, 1980b: 8-41.

Duncan, S.E. and DeAvila, E.A. "Bilingualism and Cognition: Some Recent Findings." NABE Journal 4 (1979): 15-50.

Elías-Olivares, L. and Valdés-Fallis, G. "Language Diversity in Chicano Speech Communities: Implications for Language Teaching." Sociolinguistic Working Paper No. 54, January 1979. Southwest Educational Development Laboratory, Austin, Texas.

Enkvist, N.E. "Categories of Situational Context from the Perspective of Stylistics." Survey article in Language Teaching and Linguistics: Abstracts 13 (1980): 75-94.

Erickson, J.G. and Omark, D.R. eds. Communication Assessment of the Bilingual Bicultural Child. Baltimore, Md.: University Park Press, 1981.

Ervin-Tripp, S. "An Issei Learns English." Journal of Social Issues 23 (1967):78-90.

Farhady, H. "The Disjunctive Fallacy Between Discrete Point and Integrative Tests." TESOL Quarterly 13 (1979): 347-358.

Ferguson, C.A. "Diglossia." In Language in Culture and Society, pp. 429-439. Edited by Dell Hymes. New York: Harper & Row, 1964.

_____. "Linguistic Theory." In Bilingual Education, Current Perspectives Synthesis, Vol. 5, pp. 43-52. Edited by L. Parker. Arlington, Va.: Center for Applied Linguistics, 1978.

Fishman, J.A. "Bilingualism With and Without Diglossia: Diglossia With and Without Bilingualism." Journal of Social Issues 23 (1967): 29-38.

_____; Cooper, R.L., and Ma, R. et al. Bilingualism in the Barrio. Indiana University Publications, Language Science Monograph #7. The Hague, Netherlands: Mouton & Co., 1971.

_____ and Leuders-Salmon, E. "What Has the Sociology of Language to Say to the Teacher? On Teaching the Standard Variety to Speakers of Dialectal or Sociolectal Varieties." In Functions of Language in the Classroom, pp. 77-78. Edited by C. Cazden, et al. New York: Teachers College Press, 1972.

Flavell, J.H. et al. The Development of Role Taking and Communication Skills in Children. New York: John Wiley & Sons, Inc., 1968.

García, F. et al. Multilingual Assessment Project. Final Report. Northwest Regional Educational Laboratory, Mimeo, 1975.

Gillmore, G. and Dickerson, A. "The Relationship Between Instruments Used for Identifying Children of Limited English Speaking Ability in Texas." Region IV Service Center, Houston, Texas, 1979.

Graham, C.R. and Acosta, S. "Comparison of Our State Adopted Oral Language Tests." Paper presented at the National Association of Bilingual Education, 8th Annual Conference, Seattle, Washington, 1979.

Grimm, H. "Analysis of Short-term Dialogues in 5-7 year olds: Encoding of Intentions and Modifications of Speech Acts as a Function of Negative Feedback Loops." Extended form of a paper read at the Third International Child Language Symposium, London, England, September 1975.

Gumperz, John J. "On the Linguistic Markers of Bilingual Communication." Journal of Social Issues 23 (1967): 48-57.

_____. "The Communicative Competence of Bilinguals: Some Hypotheses and Suggestions for Research." Language Society 1 (1972): 143-154.

_____ and Hernandez-Chavez, E. "Cognitive Aspects of Bilingual Communication." In Language Use and Social Change, pp. 111-125. Edited by W.H. Whitely. London: Oxford University Press, 1970.

Hansegard, N.E. Bilingualism or Semilingualism? Stockholm: Alders/Bonners, 1968.

Harris, D.P. Testing English as a Second Language. New York: McGraw-Hill, 1969.

Harris, M.L. A Comparison of Two Measures of Language Proficiency. Santa Barbara County Schools, 1979.

Haugen, E. The Norwegian Language in America: A Study of Bilingual Behavior. Philadelphia: University of Pennsylvania Press, 1953.

Helmer, S. "Demonstration of Assessment of Language Dominance of Spanish-Speaking Bilingual Children." In Occasional Papers on Linguistics: Proceedings of the First International Conference on Frontiers in Language Proficiency and Dominance Testing, pp. 226-229. Edited by James E. Redden. Southern Illinois University at Carbondale, 1977.

Herbert, C. Basic Inventory of Natural Language (BINL). San Bernardino, Ca.: Checkpoint Systems, 1977.

Hernández, L. Language of the Chicano. Bilingual Education Series, Vol. 3 (1), August 1979. National Dissemination and Assessment Center, California State, Los Angeles, 1979.

Hernández -Chávez, E. "Critique of a Critique: Issues in Language Assessment." NABE Journal 2 (1978): 47-56.

_____; Burt, M. and Dulay, H. "Language Dominance Testing: Some Behavioral Considerations." In Occasional Papers on Linguistics: Proceedings of the Second International Conference on Frontiers in Language Proficiency and Dominance Testing, pp. 137-148. Edited by James E. Redden. Southern Illinois University at Carbondale, 1978.

_____; Cohen, A. and Beltramo, A. eds. El Lenguaje de los Chicanos. Arlington, Va.: Center for Applied Linguistics, 1975.

Hess, R.D. and Shipman, V. "Early Experience and the Socialization of Cognitive Modes in Children." Child Development 36 (1965): 869-886.

Hughes, J.M. "Positive and Negative Tagging: Lexical Entries and Language Dominance in Bilinguals." In Occasional Papers on Linguistics: Proceedings of the Second International Conference on Frontiers in Language Proficiency and Dominance Testing, pp. 109-116. Edited by James E. Redden. Southern Illinois University at Carbondale, 1978.

Hymes, D. "Models of the Interaction of Language and Social Setting." Journal of Social Issues 23 (1967): 8-28. Special Issue: Problems of Bilingualism, edited by J. MacNamara, Language and Social Setting.

_____. "Introduction." In Functions of Language in the Classroom, pp. xi-iiii. Edited by C. Cazden, et al. New York: Teachers College Press, 1972.

Lanco-Worrall, A.D. "Bilingualism and Cognitive Development." Child Development 43 (1972): 1390-1400.

Ingram, E. "The Psycholinguistic Basis." In Approaches to Language Testing, pp. 1-38. Edited by Bernard Spolsky. Papers in Applied Linguistics, Advances in Language Testing, Series 2. Arlington, Va.: Center for Applied Linguistics, 1978.

Jaakkola, M. "Diglossia and Bilingualism Among Two Minorities in Sweden." International Journal of the Sociology of Language 10 (1977): 67-84.

Jakobovits, L.A. Foreign Language Learning. Rowley, Mass.: Newbury House, 1970.

_____ and Gordon, B. The Context of Foreign Language Teaching. Rowley, Mass.: Newbury House, 1974.

John, V.; Horner, V.M. and Buney, T.D. "Story Retelling: A Study of Sequential Speech in Young Children." In Basic Studies in Reading, pp. 246-262. Edited by Harry Levin and Joanna P. Williams. New York: Basic Books, 1970.

John-Steiner, V.P. and Osterreich, H. Learning Styles Among Pueblo Children. Final Report, NIE, DHEW. ED 129 503. 1975.

_____ and Nihlen, A. "The Evaluation of Communicative Competence in Bilingual Children." Bilingual Resources 2 (1979): 11-14.

Johnson, D.L. "The Influences of Social Class and Race on Language Test Performance and Spontaneous Speech of Preschool Children." Child Development 45 (1974): 517-521.

Johnson, D.M. Peer Tutoring, Social Interaction and the Acquisition of English as a Second Language by Spanish Speaking Elementary School Children. Ph.D. dissertation, Stanford University, 1980.

Jones, R.L. and Spolsky, B. eds. Testing Language Proficiency. Arlington, Va.: Center for Applied Linguistics, 1975.

Kelly, L.G. ed. Description and Measurement of Bilingualism: An International Seminar, University of Moncton, June 6-14, 1967. Canada: University of Toronto Press, 1969.

Kelly, R. "Aspects of Communicative Performance." Applied Linguistics 2 (1981): 169-179.

Kessler, C. The Acquisition of Syntax in Bilingual Children. Washington, D.C.: Georgetown University Press, 1971.

Krashen, S.A. Second Language Acquisition and Second Language Learning. Oxford: Pergamon Press, 1981.

Krauss, R.M. and Glucksberg, S. "The Development of Communication Competence as a Function of Age." Child Development 40 (1969): 255-266.

Labov, W. "The Logic of Nonstandard English." In Linguistics and the Teaching of Standard English to Speakers of Other Languages or Dialects, pp. 1-43. Edited by J. Alatis. Georgetown University Round Table on Language and Linguistics, 1969. Georgetown University, 1970.

Lado, R. Language Testing: The Construction and Use of Foreign Language Tests. New York: McGraw-Hill, 1961.

Lambert, W.E. "The Effects of Bilingualism on the Individual: Cognitive and Sociocultural Consequences." In Bilingualism: Psychological, Social, and Educational Implications, pp. 15-27. Edited by P.A. Hornby. New York: Academic Press, 1977.

_____; Havelka, W.J. and Gardner, R.C. "The Influence of Language Acquisition Contexts on Bilingualism." Journal of Abnormal and Social Psychology 56 (1958): 239-244.

Laosa, O.M. "Bilingualism in Three United States Hispanic Groups: Contextual Use of Language by Children and Adults in Their Families." Journal of Educational Psychology 67 (1975): 617-627.

Lavandera, B.R. "The Variable Component in Bilingual Performance." In International Dimensions of Bilingual Education, pp. 391-409. Edited by J. Alatis. Georgetown Round Table on Language and Linguistics, 1978. Georgetown University, 1979.

Legarreta, D. "The Effects of Program Models on Language Acquisition of Spanish Speaking Children." TESOL Quarterly 13 (1979): 521-534.

Lipski, J.M. "The Functional Relevance of Native Language in Proficiency Assessment." In Occasional Papers on Linguistics: Proceedings of the Second International Conference on Frontiers in Language Proficiency and Dominance Testing, pp. 126-136. Edited by James E. Redden. Southern Illinois University at Carbondale, 1978.

Loban, W. The Language of Elementary School Children. Research Report No. 1. Champaign, Ill.: NCTE, 1963.

Locks, N. et al. Assessment Instruments for Limited English-Speaking Students. Washington, D.C.: National Institute of Education, 1978.

Mackey, W. "The Description of Bilingualism." Journal of Canadian Linguistic Association 7 (1962): 51-85.

MacNamara, J. "The Bilingual's Linguistic Performance: A Psychological Overview." Journal of Social Issues 23 (1967): 58-77.

McClure, E. "Aspects of Code-Switching in the Discourse of Bilingual Mexican-American Children." In Linguistics and Anthropology, pp. 93-117. Edited by M. Saville-Troike. Washington, D.C.: Georgetown University Press, 1977.

Marino, B. "Early Bilingualism and Cognition: A Survey of the Literature." In Perspectives on Chicano Education. Edited by T. Gonzales and S. Gonzales. Stanford: Chicano Fellows, 1975.

Metropolitan Achievement Tests. Primary II, Form G. New York: Harcourt, Brace, Jovanovich, 1970.

Milk, R.D. Variation in Language Patterns Across Different Group Settings in Two Bilingual Second Grade Classrooms. Ph.D. dissertation, Stanford University, 1980.

Morrow, K.E. Techniques of Evaluation for a Notional Syllabus. Reading: Centre for Applied Language Studies. University of Reading, 1977.

National Institute of Education (NIE). NIE Report on the Testing and Assessment Implications of the Title VI Language Minority Proposed Rule (Internal Report). Submitted to OCR, January 12, 1981.

Office of Civil Rights (OCR). Internal Memo on LAU Enforcement Activities, FY 76-77. Education Department, 1978.

Oller, J. "Review of the Bilingual Syntax Measure." Modern Language Journal 60 (1976): 399-400.

_____. "A Program for Language Testing Research." In Papers in Second Language Research. Language Learning, special issue 4 (1976): 141-165.

_____. Language Tests at Schools: A Pragmatic Approach. London: Longman, 1979.

_____ and Hinofotis, F.B. "Two Mutually Exclusive Hypotheses About Second Language Ability: Indivisible or Partially Divisible Competence." In Research in Language Testing, pp. 13-23. Edited by John W. Oller and Kyle Perkins. Rowley, Mass.: Newbury House Publishers, 1980.

_____ and Perkins, K. Language in Education: Testing the Tests. Rowley, Mass.: Newbury House Publishers, 1978.

_____ eds. Research in Language Testing. Rowley, Mass.: Newbury House Publishers, 1980.

_____ and Richards, J. eds. Focus on the Learner: Pragmatic Perspectives for the Language Learner. Rowley, Mass.: Newbury House Publishers, 1973.

Omark, D.R. "Conceptualizations of Bilingual Children." In Communication Assessment of the Bilingual Bicultural Child. Edited by J.G. Erickson and D.R. Omark, pp. 65-67. Baltimore, Md.: University Park Press, 1981.

Overall, P.M. An Assessment of the Communicative Competence in English of Spanish Speaking Children in the Fourth and Sixth Grades. Ph.D. dissertation, Stanford University, 1978.

Palmer, A.S. "Testing Communication." IRAL 10 (1972: 35-45.

_____. "Compartmentalized and Integrated Control: An Assessment of Some Evidence for Two Kinds of Competence and Implications for the Classroom." Language Learning 29 (1979): 169-180.

Paradis, M. ed. Aspects of Bilingualism. Columbia, S. Carolina: Hornbeam Press, Inc., 1978.

Parker, L. (in consultation with Shirley Brice Heath) "Current Perspectives." In Bilingual Education, Current Perspectives: Synthesis, Vol. 5, pp. 1-62. Arlington, Va.: Center for Applied Linguistics, 1978.

Paulston, C.B. "Linguistic and Communicative Competence." TESOL Quarterly 8 (1974): 347-362.

_____. "Ethnic Relations and Bilingual Education: Accounting for Contradictory Data." In English as a Second Language in Bilingual Education, pp. 235-262. Edited by J. Alatis and K. Twaddell. Washington, D.C.: TESOL, 1976.

_____. "Viewpoint Research." In Current Perspectives in Bilingual Education: Linguistics. Arlington, Va.: Center for Applied Linguistics, 1977.

_____. "Bilingual Bicultural Education." Cited in J. Cummins, "Linguistic Interdependence and the Educational Development of Bilingual Children." Review of Educational Research 49 (1979a): 222-251,

Peak, H. "Problems of Objective Observation." In Research Methods in Behavioral Sciences, pp. 243-299. Edited by L. Festinger and D. Katz. New York: Holt, Rinehart & Winston, 1953.

Peal, E. The Effects of Bilingualism on Intelligence. M.A. thesis, McGill University, 1961.

_____ and Lambert, W.E. "The Relation of Bilingualism to Intelligence." Psychological Monographs 76 (1962): 1-23.

Peña, A.A. and Bernal, Jr., E.M. "Malpractices in Language Assessment for Hispanic Children." In Occasional Papers on Linguistics: Proceedings of the Second International Conference on Frontiers in Language Proficiency and Dominance Testing, pp. 102-108. Edited by James E. Redden. Southern Illinois University at Carbondale, 1978.

Peñalosa, F. Chicano Sociolinguistics. Rowley, Mass.: Newbury House Publishers, 1980.

Pletcher, B.P., et al. A Guide to Assessment Instruments for Limited English Speaking Students. New York: Santillana Publ. Co., 1978.

Politzer, R.L. "Linguistic and Communicative Competence, Language Dominance, Selected Pupil Characteristics and Their Relation to Achievement of Bilingual Pupils." Research Proposal to Teaching and Learning Program, NIE, 1979.

_____ and McKay, M. "A Pilot Study Concerning the Development of a Spanish/English Oral Proficiency Test." Bilingual Review 2 (1975): 112-137.

_____; Shohamy, E., and McGroarty, M. "Validation of Linguistic and Communicative Oral Language Tests for Spanish-English Bilingual Programs." Preliminary version. Presented at the Colloquium on Validation of Oral Proficiency Tests, Ann Arbor, Michigan, March 1981.

Poplack, S. Transcending Testing in Assessing Communicative Competence. Paper prepared for Language Proficiency Assessment Symposium organized by InterAmerica Research Associates, Inc., Airlie House, Virginia, March 14-18, 1981.

Ramírez, A.G. and Politzer, R.L. A Revised Spanish/English Oral Proficiency Test, 1974 Field Test Results. Research and Development Memorandum No. 134. Stanford Center for Research and Development in Teaching, Stanford University, February 1975.

Ramírez, M., et al. Spanish-English Bilingual Education in the U.S.: Current Issues, Resources and Research Priorities. Arlington, Va.: Center for Applied Linguistics, 1977.

Raven, J.E. Guide to Using the Coloured Progressive Matrices. London: H.K. Lewis & Co., Ltd., 1965.

Rodríguez-Brown, E. and Elías-Olivares, L. A Search for Congruence in Language Proficiency Testing: What the Tests Measure — What the Child Does. Paper presented at the American Educational Research Association Annual Meeting in Los Angeles, California, April 13-17, 1981.

Rosansky, E.J. "A Review of the Bilingual Syntax Measure." In Applied Linguistics, Advances in Language Testing Series 1: Some Major Tests, pp. 116-235. Edited by Bernard Spolsky. Arlington, Va.: Center for Applied Linguistics, 1979.

_____. "Issues of Validity and Reliability in Language Proficiency Assessment: A Consideration of Several Instruments." Paper delivered at the 9th Annual International Bilingual Bicultural Education Conference, April 1980, Anaheim, California.

Saer, D.J. "The Effect of Bilingualism on Intelligence." British Journal of Psychology, Vol. 14, Part 1 (1923): 25-38.

Sánchez, R. "Spanish for Native-Speakers at the University: Sugerencias." In Teaching Spanish to the Spanish Speaking: Theory and Practice, pp. 95-107. Edited by G. Valdés-Fallis and R. García-Moya. San Antonio, Texas: Trinity University, 1976.

_____. "Critique of Oral Language Assessment Instruments." NABE Journal 1 (1976): 120-127.

Senkoff, G. "Language Use in Multilingual Societies: Some Alternative Approaches." In Sociolinguistics, pp. 35-52. Edited by J.B. Pride and J. Holmes. London: Penguin Books, 1972.

SAS Institute. SAS (Statistical Analysis System) User's Guide. 1979 Edition. Cary, N.Carolina: SAS Institute, Inc., 1979.

Savignon, S. Communicative Competence: An Experiment in Foreign Language Teaching. Philadelphia: Center for Curriculum Development, 1972.

Schmidt, R.W. and Richards, J.C. "Speech Acts and Second Language Learning." Applied Linguistics 1 (1980): 129-157.

Schulz, R.A. "Discrete-point Versus Simulated Communication Testing in Foreign Languages." Modern Language Journal 5 (1977): 94-101.

_____ and Bartz, W. "Free to Communicate." In Perspective: A New Freedom, pp. 42-92. Edited by Gilbert A. Jarvis. ACTFL #7. Skokie, Ill.: National Textbook Company, 1975.

Shuy, Roger W. "Problems in Assessing Language Ability in Bilingual Education Programs." In Bilingual Education, pp. 376-380. Edited by H. LaFontain, et al. Wayne, N.J.: Avery Publishing Group, Inc., 1978.

Silverman, R.J., et al. Oral Language Tests for Bilingual Students. Northwest Regional Educational Laboratory, Portland, Oregon, 1976.

Skutnabb-Kangas, T. and Toukomaa, P. Teaching Migrant Children's Mother Tongue and Learning the Language of the Host Country in the Sociocultural Situation of the Migrant Family. Helsinki: The Finnish National Commission for UNESCO, 1976.

_____. "Semilingualism and Middle-Class Bias: A Reply to Cora Brent-Palmer." Working Papers in Bilingualism, No. 19 (October 1979): 181-198.

Southwest Regional Laboratories for Educational Research and Development. Resources for Developing a Student Placement System for Bilingual Programs. Vol. I: Oral Language, March 1980. (Office of Education, Office of Evaluation and Dissemination, Contract #300-78-0489).

Spiegel, S. and Foster, E. The "Limited-English Dominant" — How Many Are There and What Services Do They Need? Unpublished report. Los Angeles County School Resources Center, September 1980.

Spolsky, B., et al. "Three Functional Tests of Oral Proficiency." TESOL Quarterly 6 (1972): 221-235.

Stodolsky, S. and Lesser, G. "Learning Patterns in the Disadvantaged." Harvard Educational Review 37 (1967): 546-593.

Strump, T.A. "Cloze and Dictation Tasks as Predictors of Intelligence and Achievement Scores." In Language in Education: Testing the Tests, pp. 36-64. Edited by J. Oller and K. Perkins. Rowley, Mass.: Newbury House Publishers, 1978.

Thorndike, R.L. Educational Measurement. Washington, D.C.: American Council on Education, 1971.

Torrance, E.P.; Gowan, J.C.; Wu, J.M., and Aliotti, N.C. "Creative Functioning of Monolingual and Bilingual Children in Singapore." Journal of Educational Psychology 61 (1970): 72-75.

Tosi, Arturo. "Mother-Tongue Teaching for the Children of Migrants." Survey article in Language Teaching and Linguistics: Abstracts 12 (1979): 213-231.

Toukomaa, P. and Skutnabb-Kangas, T. The Intensive Teaching of the Mother Tongue to Migrant Children of Pre-school Age and Children in the Lower Level of Comprehensive Schools. Helsinki: The Finnish National Commission for UNESCO, 1977.

Tucker, G.R. "The Assessment of Bilingual and Bicultural Factors of Communication." In Bilingualism, Biculturalism and Education pp. 217-240. Edited by S.T. Carey. Edmonton: University of Alberta Press, 1974.

Ulibarri, D.M. Technical Manual: Multilingual Assessment Scale. Multilingual Assessment Program: Stockton Unified School District, Stockton, California, 1974 (USOE; Title VII).

_____. Cognitive Processing Theory and Culture-Loading: A Neo-Piagetian Approach to Test Bias. Ph.D. dissertation, University of California, Berkeley (in preparation).

_____; Spencer, M.L. and Rivas, G.A. "Comparability of Three Oral Language Proficiency Instruments and Their Relationship to Achievement Variables." NABE Journal 5 (1981): 47-80.

Valette, R.M. Modern Language Testing. New York: Harcourt, Brace, Jovanovich, Inc., 1967.

_____. Modern Language Testing, 2nd edition. New York: Harcourt, Brace, Jovanovich, Inc., 1977.

Vernon, P.E. ed. Secondary School Selection: A British Psychological Society Inquiry. London: British Psychological Society, 1957.

Walters, J. Language Variation in the Assessment of the Communicative Competence of Bilingual Children: Evidence for the Linguistic Interdependence Hypothesis. National Dissemination and Assessment Center, California State, Los Angeles, 1979.

Weinreich, U. Languages in Contact. Originally published as Number 1 in the series "Publications of the Linguistic Circle of New York, 1953." The Hague: Mouton & Co., 1974.

Wiemann, J.M. and Backlund, P. "Current Theory and Research in Communicative Competence." Review of Educational Research 50 (1980): 185-199.

Zentella, A.C. Code-switching and Interactions Among Puerto Rican Children. Working Papers in Sociolinguistics, No. 50, 1979.

Zirkal, P. "A Method for Determining and Depicting Language Dominance." In English as a Second Language in Bilingual Education pp. 133-143. Edited by J. Alatis & K. Twaddell. Washington, D.C.: TESOL, 1976.

APPENDICES

DECISION RULES FOR SAMPLE #2 SELECTION

Score Possibilities

HIGH SPANISH/HIGH ENGLISH	N
LAB scores ≥ 65% in English and Spanish	1
LAB scores ≥ levels 4 or 5 in English and Spanish	10
BSM scores ≥ levels 4 or 5 in English and Spanish	1
Fluent English/Fluent Spanish (previous data, no other information)	0
High Reading Spanish/Fluent English	15
Any combination of the above	9
1977-79 LAB scores Spanish ≥ level 4 or 5/English ≥ level 3	1
1977-78 LAB scores Spanish ≥ level 4 or 5/English ≥ level 3	1
1977-78 BSM scores Spanish ≥ level 4 or 5/English ≥ level 3	1
Other	9

LOW SPANISH/LOW ENGLISH	N
LAB scores ≤ 65% in English and Spanish	10
LAB scores ≤ level 3 in English and Spanish	3
BSM scores ≤ level 3 in English and Spanish	0
LIM-NCM English/LIM-NCM Spanish (previous data, no other information)	0
Low Spanish reading/LIM-NCM English	13
Any combination of the above	4
1978-79 Low Spanish/Low English	1
Other	1

MISCELLANEOUS	N
NO DATA	10
English Only Data	39
Spanish Only Data	21
Previous scores - High Spanish/Low English (1976-78)	19
Current scores - LAB English 65%/LAB Spanish 40%-61%	3
Previous scores - Low Spanish/Low English (1976-78)	1

DISTRIBUTION OF SAMPLE #2

LO LO	N	HI HI	N
Low Spanish/Low English	45	High Spanish/High English	52
Spanish Only Data - Low Spanish	9	Spanish Only Data - High Spanish	11
English Only Data - Low English	19	English Only Data - High English	20
Low Spanish/Average English	4	Previous High Spanish/Low English	19
Totals	77	Totals	102

LO LO	77
HI HI	102
NO DATA (Random Selection)	6
Total in Sample #2	185

SAMPLE #2 CROSSTABULATIONS
OF SPANISH BSM II LEVEL BY ENGLISH LEVEL

N=176

BSM ENGLISH LEVEL

BSM SPANISH LEVELS	1	2	3	4	5	6
1	1	0	0	0	11	3
2	0	0	0	0	0	1
3	0	0	0	5	19	1
4	0	0	1	2	13	4
5	0	1	2	12	41	4
6	3	0	5	9	36	2

TEACHER COMMUNICATIVE COMPETENCE RATING SCALE

1. Student communicates quite effectively in English/Spanish; almost always gets meaning across.

2. Student communicates fairly effectively in English/Spanish; generally gets meaning across.

3. Student has difficulty communicating effectively in English/Spanish; sometimes gets meaning across.

4. Student does not communicate effectively in English/Spanish; generally does not get meaning across.

5. English/Spanish communication skills are virtually absent; student cannot be tested in English/Spanish.

PARENT PERMISSION SLIP

(English)

December 10, 1980

Dear Parents,

Ms. Zoe Ann Hayes, from Stanford University, will be conducting a research project in our school. The purpose of the project is to describe the language of children who are bilingual. Your child, _____, has been chosen to participate in this project because he/she has the very special ability of knowing two languages.

Ms. Hayes has been involved with bilingual children as a teacher's aide and bilingual teacher. She is currently a professor of bilingual education; this project is for her dissertation at Stanford University. The project has been funded by InterAmerica Research Associates of Rosslyn, Virginia, as one of several similar studies throughout the United States.

With your consent _____ will participate in language activity sessions using puppets, cartoons, games and pictures. The sessions will be tape recorded for analysis of language. The sessions will be in both English and Spanish. _____ will go to a room for each session to work individually with a bilingual assistant. One group session will also be needed. _____ may go to 8 sessions in all. The project will last from January to May. If you will be leaving the area before June 1981, please fill in the date you plan to leave on the form at the bottom of this letter.

You will be asked to provide background information. Ms. Hayes will need to know how long you and your child have been living in the United States and a few other details. You will be contacted by a bilingual assistant in March for this information.

All information about you and your child will be confidential. Your names will never be used. No school personnel nor outside agents will have access to individual student information. The information will only be used for this project.

If you are not satisfied with any part of the study you may withdraw your child at any time. Also, you may call Stanford University, Sponsored Projects Office (415-497-2883) to report your complaints anonymously.

I believe this project will help us to develop better educational programs for bilingual students like _____. If you are willing for your child to participate in this study, please sign this form and mail it in the attached addressed, stamped envelope.

Thank you.

_____

Teacher

I give permission to have my child, _____, participate in Ms. Hayes' project.

_____

Signature of parent or guardian / date

We plan to move from the area in the month of _____.

PARENT PERMISSION SLIP

(Spanish)

10 de diciembre 1980

Estimados Padres,

La señorita Zoe Ann Hayes de la Universidad de Stanford va a hacer un proyecto en nuestra escuela. El propósito del proyecto es de describir el lenguaje de los niños bilingües. Su hijo(a) _____ ha sido escogido(a) para participar en este proyecto porque él (ella) tiene esa habilidad tan especial de hablar dos idiomas.

La señorita Hayes ha trabajado como ayudante bilingüe al profesor y como maestra bilingüe. Ahorita trabaja como profesora de educación bilingüe. Este proyecto es para su tesis de Stanford. El proyecto ha conseguido fondos de la organización InterAmerica Research Associates (Asociados para Investigaciones Interamericanas) basada en Roselyn, Virginia. Este proyecto es uno de varios parecidos que se llevan a cabo en diferentes partes de los Estados Unidos.

Con su permiso _____ participará en algunas actividades que incluirán caricaturas, juegos, dibujos y títeres o muñecos. El lenguaje de todos los participantes en las actividades se gravará en una grabadora para luego analizarlo. Las actividades se llevarán a cabo en inglés y en español. _____ irá a un cuarto especial para cada sesión o actividad donde trabajará individualmente con un asistente bilingüe. Además una vez participará en una sesión con un grupo de estudiantes. En total puede ser que _____ participará en 8 sesiones.

El proyecto se realiza desde enero hasta mayo. Si su familia tiene planes de irse de esta área antes de junio de 1981, por favor díganos la fecha cuando piensan salir. Hay un lugar en la forma incluida para decirnos esta información.

La señorita Hayes necesitará un poco de información de usted. Por ejemplo, ¿cuántos años usted y su hijo(a) tienen viviendo en este país ? En marzo un asistente bilingüe les llamará para conseguir esta información.

Toda la información y las gravaciones serán confidenciales. Nunca usará la señorita Hayes el nombre de usted o de su hijo(a) en sus reportages. No podrán ver la información ni personal de la escuela, ni gente de afuera, ni organizaciones. La información se usará solamente para este proyecto.

Si a usted no le parece bien alguna actividad, pregunta o parte del proyecto puede quitar a su hijo(a) del proyecto en cualquier momento. También si desea, usted puede llamarle a la Oficina de Proyectos (Sponsored Projects Office) al número 415-497-2883 para reportar, sin dar su nombre, cualquier incidente o queja que tenga.

Yo creo que este proyecto nos va a ayudar a desarrollar mejores programas educacionales para nuestros hijos bilingües como _____. Si usted está de acuerdo que su hijo(a) participe en este proyecto, favor de firmar la forma incluida abajo y mandarla por correo en el sobre incluido con timbre.

Gracias.

_____

Maestro/ Maestra

Le doy permiso a mi hijo(a) _____ de participar en el proyecto de la señorita Hayes.

_____

Firma del padre, madre o guardián / fecha

Tenemos planes de mudar o salir de esta área en el mes de _____.

BSM II Level Descriptions (English and Spanish)*

Level	Description
Level 1: No English	Except for a word here or there, students at this level are able to understand little or no English and are not able to speak it at all.
Level 2: Receptive English only	Students at this level are able to understand conversational English in varying degrees. Also, these students are able to produce some Common English words and phrases spontaneously and repeat short sentences or questions, but they are unable to use English as a vehicle for significant communication.
Level 3: Survival English	Level 3 students can usually make themselves understood by using a combination of simple speech, gestures, and an occasional word from their native language. When speaking English, these students sometimes omit nouns or verbs, replacing them with gestures or words from their native language. In addition, they make many errors in the use of articles, verb endings and pronouns. These students are usually able to communicate ideas and feelings in English, but only with considerable difficulty.
Level 4: Intermediate English	Students at Level 4 have little difficulty communicating their ideas in English, and do not rely heavily on gestures or on their native language to get across their meaning. They usually control syntactic structures that include plurals, articles, pronouns, and some verb endings. Errors are often made in the more complex verb forms, and in advanced number agreements.
Level 5: Proficient English I	Level 5 students demonstrate a fairly high degree of proficiency in English, approaching native proficiency in the case of younger students. They control most of the basic grammatical structures of English. For older students, this level represents incomplete learning of some of the more advanced structures.
Level 6: Proficient English II	Students at Level 6M have mastered a broad range of the syntactic structures found in the speech of native speakers of English through high school age. Certain nonstandard forms are used. Students at Level 6S have mastered a broad range of the standard syntactic structures found in the speech of native speakers of English through high school age. Levels 6M and 6S are intended to represent comparable degrees of syntactic proficiency in English. However, depending upon specific program goals, the 6S/6M distinction may be given additional weight.

*For Spanish descriptions, replace "English" with "Spanish."

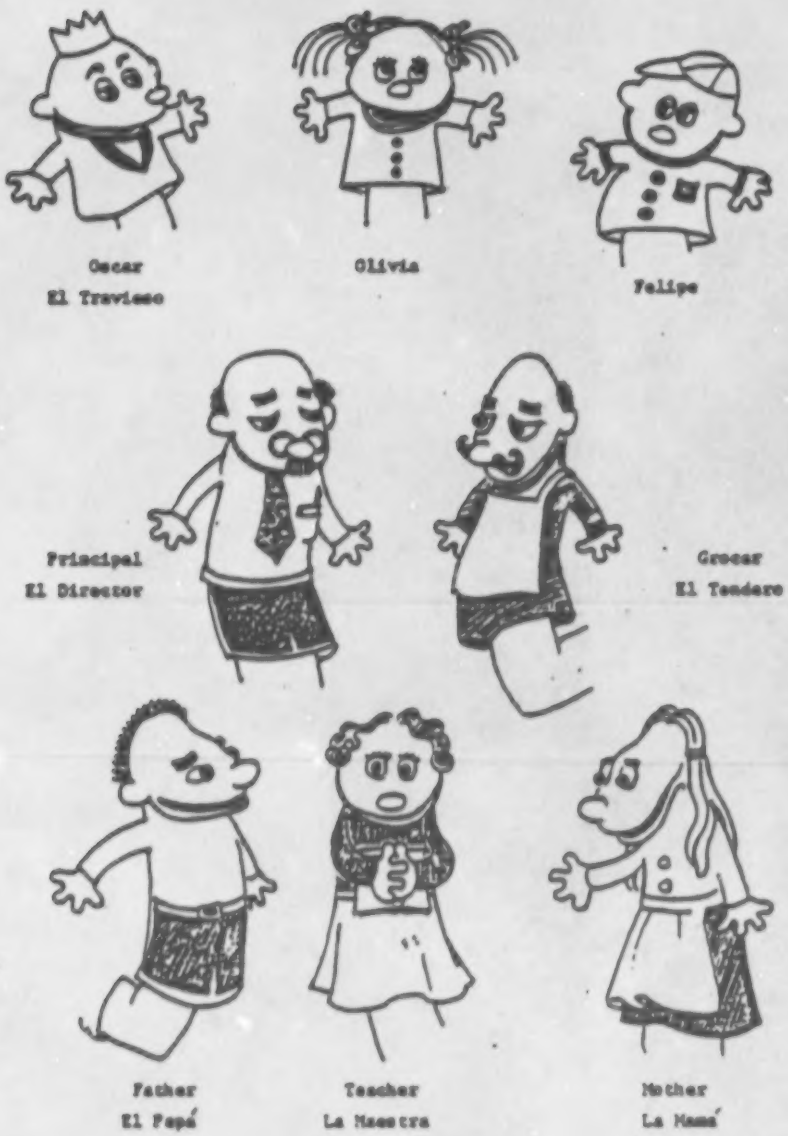
Source: Burt, Dulay et al., 1980.

SELECTED EXAMPLES OF BSM II ITEMS (in response to pictures)

Language	Item #	Question	Skill
English	11	What would happen if the good family came back now?	conditional
Spanish	18	¿Si regresara ahora su familia que haría ella?	conditional
English	20	They don't know what happened, do they? Why is the food gone?	irregular past
Spanish	3	¿Qué hicieron los niños con los zapatos cuando se los quitaron?	irregular past

SOURCE: Burt, M.K., Dulay, H.C. and Hernández-Chávez, E. Bilingual Syntax Measure II (Medida de Sintaxis Bilingüe II): Student Response Booklet (Folleto de Respuestas del Alumno). New York: Harcourt, Brace, Jovanovich, Inc., 1978,1980.

CCT Puppets


Oscar
El Travieso
Olivia
Felipe
Principal
El Director
Grocer
El Tendero
Father
El Papá
Teacher
La Maestra
Mother
La Mamá

COMMUNICATIVE COMPETENCE TEST

Puppet Interaction Situations

1. Felipe* wants his father to buy him a bicycle for his birthday. You be Felipe.

2. Felipe was fighting in the classroom. He has to go to the principal's office. You be Felipe. (Overall, 1978)

3. Felipe and Olivia are brother and sister. They are watching television. Felipe wants to watch Batman, and Olivia wants to watch cartoons. You be Felipe.

4. Felipe was late for school. He has to tell the teacher why he was late. You be Felipe. (Overall, 1978)

5. The mother wants Felipe to go to buy some tortillas at the store. You be Felipe.

6. This is Oscar. He is very stubborn and mean. He always wants to bother other people, and he never stops. He is going to scribble all over Felipe's work paper. You be Felipe.

EXTRA SITUATIONS

(if the child does not answer 1, 3 or 5, use this item) Felipe wants to spend the night at his friend, Jaime's house. He has to ask his father for permission to sleep over.

(if 2, 4 or 6 are not answered, use this item) The teacher wants Felipe to go to the office for some writing paper. You be Felipe.

*If a girl is being tested, a girl puppet named Olivia is used.

EXAMPLE OF TESTER SCRIPT FOR PUPPET INTERACTION SITUATIONS: CCT

Tortillas

(English)

The mother wants Felipe to go to buy some tortillas at the store. You be Felipe.

Tester: (as mother) Felipe, please hurry to the store and buy some tortillas for dinner. Here's a quarter.

Student:

Tester: (as grocer) Oh, Felipe, how are you?

Student:

Tester: How can I help you?

Student:

Tester: I'm sorry, but there aren't any more tortillas. We have some delicious bread. Would you like some bread instead?

Student:

Tester: (if the child says no) OK, sorry. Say hello to your mother for me.

(if the child says yes) OK, here you are. That will be a quarter. Bye.

Tester: (as mother) Felipe, just in time. Where are the tortillas?

Student:

Tester: (take it naturally to a positive close from here)

EXAMPLE OF TESTER SCRIPT FOR PUPPET INTERACTION SITUATIONS: CCT

Tortillas

(Spanish)

La mamá mandó a su hijo, Felipe, a comprar tortillas a la tienda. Haz el papel de Felipe.

Tester: (mamá) Felipe, vete a la tienda a comprar tortillas para la cena. Date prisa. Aquí tienes una peseta.

Student:

Tester: (tendero) O, Felipe, ¿Cómo estás?

Student:

Tester: ¿En qué te puedo servir?

Student:

Tester: Ok, fíjate, se acabaron las tortillas. Pero tenemos pan muy bueno. ¿No quieres pan?

Student:

Tester: (si dice no) OK. Saludos a tu mamá.

(si dice sí) OK. Aquí tienes. Una peseta, por favor.

Tester: (mamá) O, Felipe, ¡por fin! ¿Dónde están las tortillas?

Student:

Tester: (take it naturally to a positive close from here)

EXAMPLE OF TEST TRANSCRIPT: CCT

Tortillas

(actual speech)

Tester: "La mamá quiere que Felipe compre unas tortillas en la tienda, ¿OK?"

Mother: "Felipe! Toma esta peseta y cómprame un paquete de tortillas. ¡Y rápido, rápido!"

Felipe: "No, porque no cuestan una quarter."

Mother: "¿ Cuánto cuestan?"

Felipe: "Un dólar."

Mother: " ¡No, mentiroso! Aquí tienes cincuenta centavos."

Felipe: "No me alcanza."

Mother: " ¿No? Traígame el cambio, ¿OK? Aquí tienes un dólar."

Grocer: "Hola, Felipe. ¿Cómo está? ¿En qué le puedo servir?"

Felipe: "Unas tortillas."

Grocer: "Se acabaron las tortillas. ¿Quiere un poco de pan?"

Felipe: "Sí."

Grocer: "OK. Un dólar, por favor. Gracias. Saludos a tu mamá."

Mother: "¡Ah Felipe, ya llegas con las tortillas!"

Felipe: "No había."

Mother: "Entonces, ¿qué?"

Felipe: "Compré pan."

Mother: "Está bien. ¿Y el cambio?"

Felipe: "No sobró nada."

Mother: "OK. Gracias Felipe."

SCHOOL STORY SEQUENCE SCENE: CCT

HERE IS A LONGER STORY. LOOK AT THESE PICTURES. THEN TELL ME THE STORY.
(Show the child the school scene. Wait 10 Seconds. Keep the tape recorder on!!)
[AQUI HAY UN CUENTO UN POQUITO MAS LARGO. MIRA LOS DIBUJOS Y LUEGO DIME LA HISTORIA.]

GOOD. HERE THE TEACHER IS PROBABLY SAYING: "Who knows the answer to this math problem?" (point)

WHAT DO YOU THINK SHE IS SAYING HERE? (point)

WHAT DO YOU THINK THE BOY'S FRIENDS ARE SAYING? (point)

WHAT DOES THE NURSE SAY? (point)

WHAT DOES THE PRINCIPAL SAY? (point)

AND WHAT DOES THE PRINCIPAL SAY TO THE MOTHER? (point)

[MUY BIEN. AQUI LA MAESTRA A LO MEJOR DICE: "¿Quién sabe la respuesta a este problema?"

¿QUE CREES QUE DICE AQUI?

¿QUE CREES QUE LE DICEN SUS COMPANEROS?

Y ¿QUE LE DICE LA ENFERMERA?

¿QUE LE DICE EL DIRECTOR?

Y ¿QUE LE DICE EL DIRECTOR A SU MAMA?]

GLOBAL SPEAKING PROFICIENCY SCALE

1. Speaker has almost no proficiency. Virtually incomprehensible speech; few vocabulary items; grammar weak.

2. Speaker has minimal proficiency. Speech is hard to understand, but speaker has somewhat larger vocabulary than Level 1; grammar poor, but used in a few acceptable structures at least; speaker can express a few ideas, but lack of fluency impedes understanding.

3. Speaker has fair proficiency. Speech is largely understandable; basic vocabulary items and grammatical structures controlled; moderate fluency with some stretches of hesitation and false starts.

4. Speaker has good proficiency. Speech is understandable; all necessary vocabulary items and grammatical structures controlled, and some more elaborate patterns used correctly; fluency good, few if any hesitations.

5. Speaker has excellent proficiency. Speech is entirely fluent and comprehensible; extensive vocabulary and grammar controlled; correct idiomatic use of English/Spanish expressions present.

SOURCE: Politzer, R. Working manuscript. CERAS. Stanford, Ca., 1981.

GRAMMATICALITY SCALE

1. Student's use of grammatical structures is very limited; usually responds in single word utterances or with "Yes", "No", or "I don't know".

2. Student attempts simple grammatical structures; these are characterized by many errors.

3. Student attempts some complex grammatical structures; these are characterized by many errors; uses simple structures with some errors.

4. Student's use of complex structures is better than Level 3; some errors occur; uses simple structures without errors.

5. Student uses complex grammatical structures with very few errors.

VOCABULARY PROFICIENCY SCALE

1. Speaker has very limited vocabulary, consisting of the names of a few nouns in the pictures; there are no or few verbs or prepositions used.

2. Speaker has somewhat larger vocabulary than in Level 1; vocabulary still limited generally to nouns, but also includes some verbs and prepositions. Speaker may use some phrases; there is considerable translation of the L1 into English; this may produce non-native forms.

3. Speaker's vocabulary includes most nouns in the tasks, common verbs and some of their correct compound forms (i.e., pick up). The circumlocutions used are rather general or vague, such as "the man in the truck" for "ice cream man" or "small animal" for "bird". Some translation from L1.

4. Speaker's vocabulary is extensive and fairly specific for all tasks. Many prepositions used correctly. The circumlocutions which may occur are quite precise.

5. Vocabulary is accurate, precise, totally appropriate to speaker's age and developmental level. Speaker has command of idiomatic expressions.

SOURCE: Politzer, R. Working manuscript. CERAS. Stanford, Ca., 1981.

COMPREHENSION SCALE

1. Student understands too little for the simplest type of conversation.

2. Student understands only slow, very simple speech on common topics; requires constant repetition and rephrasing.

3. Student understands careful, somewhat simplified speech directed to him or her, with considerable repetition and rephrasing.

4. Student understands quite well normal speech directed to him or her, but requires occasional repetition or rephrasing.

5. Student understands everything in normal conversation, except for very colloquial or low-frequency items or exceptionally rapid or slurred speech.

SOURCE: Adapted from FSI numerical rating procedure. Valette, 1977: 159.

FLUENCY RATING SCALE

1. So halting and fragmentary that conversation is virtually impossible. Inefficient use of time in conveying information.

2. Speech is frequently jerky and hesitant. Utterances and sentances left uncompleted. Minimally efficient use of time in conveying information.

3. Speech is occasionally hesitant and jerky. Some unevenness caused by rephrasing and groping for words. Somewhat efficient use of time in conveying information. (Lack of efficiency due to circumlocution, rephrasing, etc.)

4. Speech is usually effortless and smooth, but perceptibly non-native in speed and evenness. Efficient use of time in conveying information.

5. Always native-like in speed and evenness. Optimally efficient use of time in conveying information.

SOURCE: Politzer, R. Working manuscript. CERAS. Stanford, Ca., 1981.

AMOUNT OF RELEVANT COMMUNICATION SCALE

1. Virtually no relevant information was conveyed by the student. (0 - 20%)

2. Very little relevant information was conveyed by the student. (25%)

3. Some relevant information was conveyed by the student. (50%)

4. Most relevant information was conveyed by the student. (75%)

5. All relevant information was conveyed by the student. (90 - 100%)

SOURCE: Adapted from Bartz, 1974.

EFFORT TO COMMUNICATE SCALE

1. Student withdraws into long periods of silence, without any apparent effort to complete the tasks (small amount of effort).

2. Student makes little effort to communicate; what he or she does do is "half-hearted", without any enthusiasm.

3. Student makes some effort to communicate, but still shows a somewhat "disinterested" attitude.

4. Student makes a real effort to communicate; attempts to answer all questions.

5. Student makes a special (unusually high) effort to communicate.

SOURCE: Adapted from Bartz, 1974.

WORD NAMING BY DOMAIN

ID Number _____

VOCABULARY

Trial Item

DIRECTIONS

KITCHEN: Now let's look at these pictures. This is a kitchen. Look at all of the things in the kitchen. Let's see how many things we can name: (TESTER - point to objects and name them).

Can you name some more?

What else do you see?

Is that all? GOOD

HOME: Now let's look at these pictures. They are pictures of the rooms in a house. Here is the bathroom, the bedroom, the living room and the hall. I want you to name all of the things you can in these pictures. I am going to write down what you say (TESTER- write down words on HOME page).

Can you name some more?

What else do you see?

Is that all? GOOD

SCHOOL: Now let's look at this school room. See all of the things in the room? I want you to name as many things as you can. I am going to write down what you say.(TESTER- write down words on SCHOOL page).

Can you name some more?

What else do you see?

Is that all? GOOD

THANK YOU VERY MUCH . YOU DID VERY WELL.

Source: The First Thousand Words by H. Amery and illustrated by Stephen Cartwright. New York: Mayflower Books, Inc. 1979.

BEST COPY AVAILABLE

ID. Número

VOCABULARIO

DIRECCIONES

COCINA: Ahora vamos a ver estos dibujos. Este es la cocina. Mire a todas las cosas en la cocina. Vamos a ver cuantas cosas de la cocina podemos nombrar. (TESTER- start naming objects and pointing to them).

¿Puedes nombrar más?

¿Y qué más?

¿Es todo? BUENO

CASA: Ahora estos dibujos son de la casa. Aquí hay el baño, la recámara, la sala y el pasillo. Quiero que nombres lo que puedas de las cosas en la casa. Voy a puntar lo que me dices. (TESTER- write down words on CASA page).

¿Puedes nombrar más?

¿Y qué más?

¿Es todo? BUENO

ESCUELA: Ahora vamos la clase de la escuela. ¿Ves las cosas en la clase? Quiero que nombres lo que puedas de las cosas en la clase. Voy a puntar lo que me dices. (TESTER- write down words on ESCUELA page).

¿Puedes nombrar más?

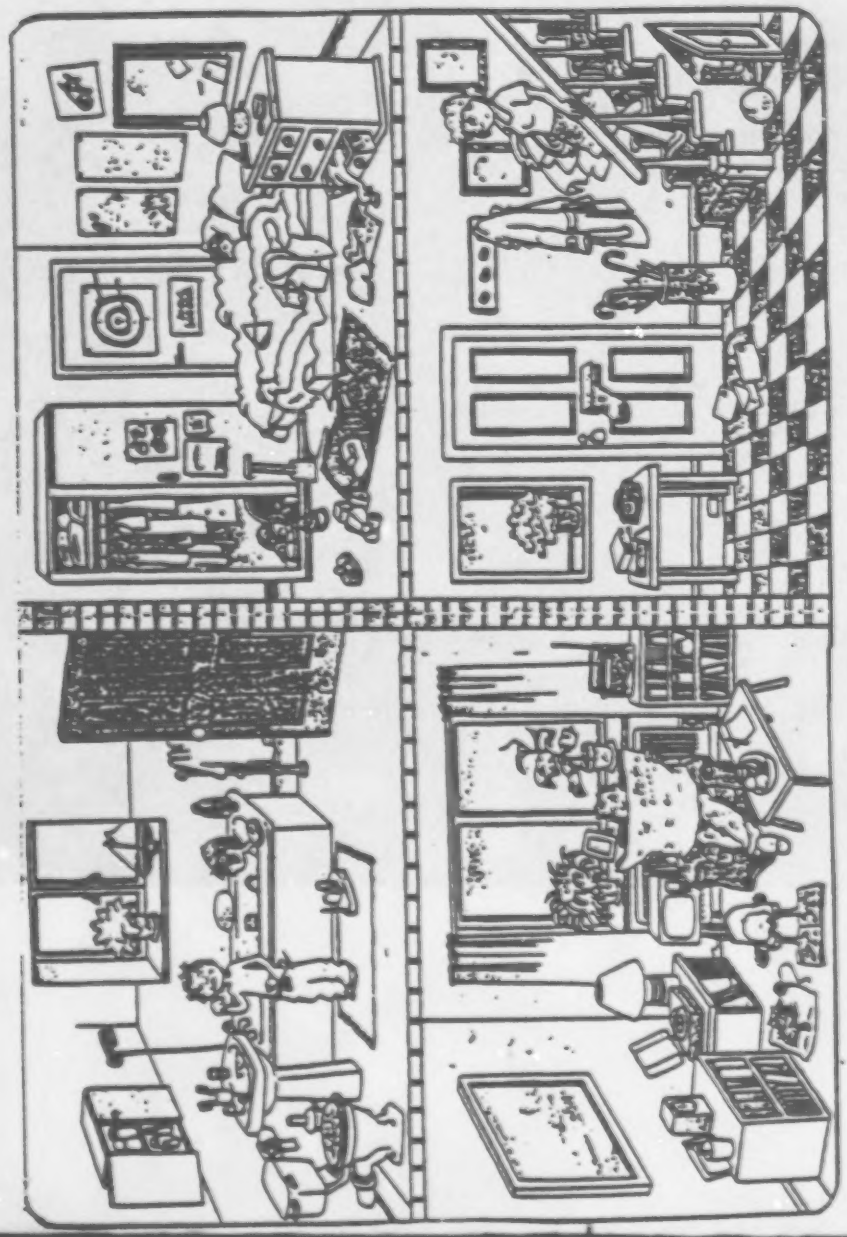
¿Y que más?

¿Es todo? BUENO

BEST COPY AVAILABLE

Muchas Gracias. ¡Lo hiciste muy bien!

WNBD: Home Picture



WMED: School Picture

EXAMPLES FROM THE CERAS BALANCE

Administrator's Manual

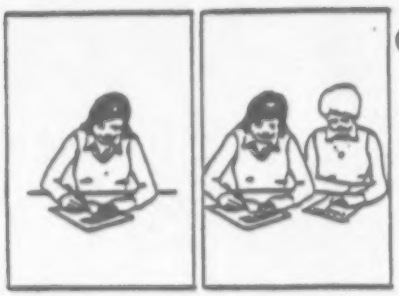
(English)

THE CHILD IS SINGING.

A. THE CHILDREN SINGING.

B. THE CHILDREN IS SINGING.

C. THE CHILDREN ARE SINGING.



THIS WOMAN IS WRITING.

A. THESE WOMEN ARE WRITING.

B. THESE WOMEN WRITTEN.

C. THESE WOMEN ARE WRITTEN.

THIS CHILD IS EATING.

A. THESE CHILDREN EATS.

B. THESE CHILDREN EATEN.

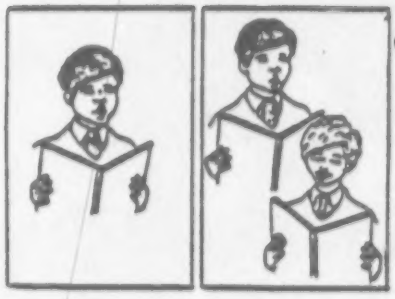
C. THESE CHILDREN ARE EATING.

Source: Adapted from Ramírez & Politzer, 1975.

EXAMPLES FROM THE CERAS BALANCE

Administrator's Manual

(Spanish)



1. EL NIÑO ESTÁ CANTANDO.
   - A. LOS NIÑOS CANTANDO.
   - B. LOS NIÑOS ESTÁ CANTANDO.
   - C. LOS NIÑOS ESTÁN CANTANDO.

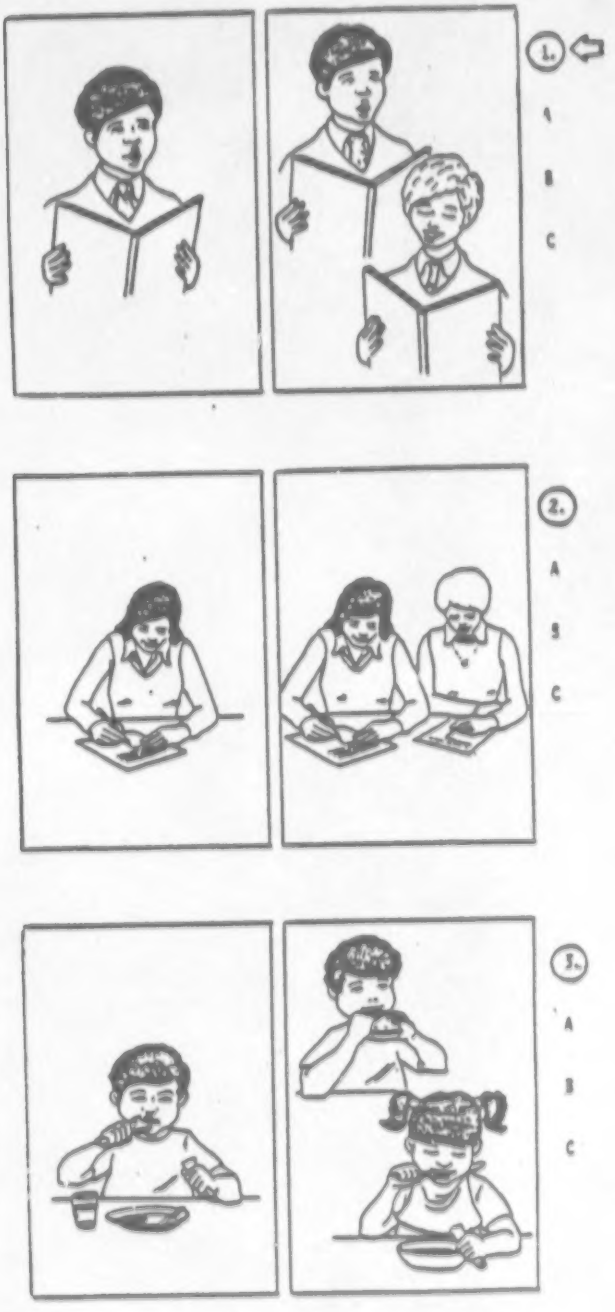
2. ESTA MUJER ESTA ESCRIBIENDO.
   - A. ESTAS MUJERES ESTÁ ESCRIBIENDO.
   - B. ESTAS MUJERES ESTÁN ESCRIBIENDO.
   - C. ESTAS MUJERES ESTÁ ESCRITO.

3. ESTE NIÑO ESTÁ COMIENDO.
   - A. ESTOS NIÑOS ESTÁN COMIENDO.
   - B. ESTOS NIÑOS ESTUVIERON COMIENDO.
   - C. ESTOS NIÑOS ESTÁN COMIÉNDOSE.

Source: Adapted from Ramírez & Politzer, 1975.

EXAMPLES FROM THE CERAS BALANCE

Student Response Booklet

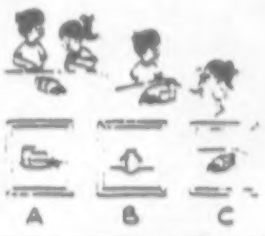


Source: Adapted from Ramírez & Politzer, 1975.

EXAMPLE FROM THE CCS

[illegible]: 1,13,20,25

These items involve selecting a picture which shows what a group of objects looks like from a perspective other than the child's own.



DO make sure story is understood.

DO indicate the position from where girl is taking picture.

Script (Item 1)

Say	Do
See the camera and the toys? Well, the boy wants the girl to take a picture of the toys.	Point to the camera and the toys, and then to the girl.
What if she takes the picture from over here?	Point to where the girl is in the second scene.
See where she's taking the picture from?	Again point to indicate where she is taking the picture from.
Well, what do you think the picture will look like if she takes the picture from over here?	Point to where the picture is being taken from.
Do you think the picture will look like this one?	Point to the first answer (A).
Or do you think it will look like this one?	Point to the second answer (B).
Or do you think it will look like this one?	Point to the third answer (C).
O.K., mark an "X" on the one you think makes the story true.	Be sure the story is understood.
Does everyone have an "X"? Good, let's go to the next story.	Check to see that each child has marked an "X."

Source: DeAvila, 1980.

BEST COPY AVAILABLE